

Bringing enlightenment to the pursuit of justice

Advances in modern science and medicine have introduced unprecedented dilemmas for both courts and legislators. A ruling in a British case on artificial insemination gives hope that good sense can be made to prevail.

Last week, a British appeals court overruled a decision by the Human Fertilization and Embryo Authority (HFEA) preventing a woman from travelling to Belgium to be inseminated with the sperm of her dead husband. The verdict, directing the HFEA to reconsider its decision, is the latest step in a complex legal battle arising from the fact that the husband had not given his written permission for the insemination. Coincidentally, the verdict was announced on the day that a brief meeting opened in Paris of leading judges and legal experts from around the world, called to debate the rapidly growing interactions between science, medicine and the law.

Until recently, both science and medicine were carried out relatively free of legal constraints (apart from considerations of public safety). One point that became clear at the meeting is that both are increasingly affected by an arsenal of regulations designed, for example, to control the applications of recombinant DNA technology and medically assisted procreation. But there was some disarray among those present about how the law should be applied — if at all — to such complex and fast-moving areas. In particular, doubts were expressed about the legitimacy of using the law to rule on matters that depend partly on the morality of a certain culture and partly on the conscience of individual scientists, physicians and patients.

Some of those present, particularly from France and elsewhere in Europe, argued that the guardians of national constitutions, such as supreme courts and constitutional councils, should be recognized as the ultimate arbiters of bioethical issues, particularly given the growing proliferation of situations in which these issues arise. They claim that a reaffirmation of the principles enshrined in a national consti-

tution is necessary in such situations to defend fundamental concepts of social justice and individual liberty, including scientific freedom, the non-commercial value of the human body, the need for informed consent to human experiments and — a growing area of concern — the protection of 'genetically disadvantaged' individuals.

But others say that constitutional elements may be relevant only in a few areas. Stephen Breyer, for example, a judge of the US Supreme Court, told last week's meeting, 'Constitution and Bio-medical Ethics', that many bioethical issues can be adequately dealt with by administrative law, by penal law, or by the decisions of individual ethics committees. Or they can simply be left to market forces. At the same time, Breyer admits that making ideas about the basic rights of human beings explicit in constitutions and international treaties can highlight important considerations and make legislators "stop and think" when introducing new laws.

Last week's case in Britain vividly illustrates the type of dilemma faced by modern courts. The case arose largely because the husband of the woman concerned died unexpectedly — of meningitis — before he was able to give his written consent, a situation that those who drafted the legislation setting up the HFEA admit candidly that they had not foreseen. Some creative legal footwork, involving the case being characterized as exceptional, has been necessary to stay, at least technically, within this legislation. But, for the present, it seems human justice has prevailed. If the various legal systems in place across the world can apply the same measure of principled creativity to similar issues elsewhere, there is no reason why carefully crafted regulations should impede either scientific or human advancement. □

Heavenly prospects and glad tidings

President Clinton has accepted that science should not suffer from his efforts to balance the budget.

After two years of relentless gloom about their long-term funding prospects, scientists in the United States have cause this week for some mild celebration. There is good news in Japan, too, where a very different set of economic circumstances is leading to greater government investment in science.

President Bill Clinton's 1998 budget offers steady funding for most science agencies, and, in its projections for future years, lifts the threat that very deep cuts would have to take place in order to balance the US budget by 2002 (see page 565). The budget scarcely qualifies as a windfall for science; most agencies would obtain increased funding close to the rate of inflation. But, given the dark expectations of a year ago, it is being welcomed by a community that now holds high hopes that the Congress will ramp up the proposed increases.

The main reason for the improved science funding outlook is extraneous: Clinton has taken advantage of both improved economic projections and the political disarray of his opponents to present a budget-balancing plan that includes few real cuts in non-defence government spending. In practice, his budget will increase the deficit by \$5 billion

next year, at a time when rapid economic growth ought to allow it to fall. But science lobbyists also detect, with some justification, evidence that their message has finally been getting through in Washington. The president himself has certainly shown increased interest in science over the past six months. The change appears to date from his successful television appearances praising NASA's alleged discovery of life on a meteorite from Mars. In a series of interviews, and subsequently on the campaign trail last autumn, the president began to warm to scientific themes (see *Nature* 383, 566; 1996). By the time of his fourth State of the Union address last week — his first to mention science at all — he was ready to embark upon a lengthy eulogy, stressing space exploration, AIDS research and the Human Genome Project.

Not to be outdone, Republicans in the Senate have come up with their own bold initiatives to boost spending on science. Jack Gibbons, Clinton's science adviser, has observed that "the public and the vast majority of the Congress support a strong research program". As long as the public and the Congress are willing to join the president in turning a blind eye to the budget deficit, it appears that they will get one. □



Outlook brightens for science in US budget proposals for 1998

[WASHINGTON] President Bill Clinton has proposed increases close to the rate of inflation for most US science agencies in his 1998 budget, lifting a threat that his plans to balance the overall federal budget by 2002 would severely hit research funding.

The budget plan for the next financial year, which begins in October, would increase total federal spending on research and development (R&D) by 2 per cent, to \$75.5 billion. Non-defence R&D would grow by 4 per cent to \$35 billion. The proposal will now be considered by the Congress, which is due to pass final budget bills by 1 October.

Advocates for science gave the budget a cautious welcome, pointing out that the spending proposed for 1998 comfortably exceeds that projected by the administration 12 months ago. And new projections to 2002 remove the threat of sharp cuts at the National Aeronautics and Space Administration (NASA), the National Science Foundation (NSF) and the Department of Energy.

"Given the tenor of the times, this is quite an optimistic budget," says Cornelius Pings, president of the Association of American Universities. Pings says that the budget proposals suggest that "the concerted efforts of the science community to engage the administration are perhaps beginning to pay off".

At the same time, there was criticism of the president for failing to assure real growth in the 1998 budgets of the NSF and the National Institutes of Health, the two agencies that fund most university science.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Clinton: proposals would avoid major cuts but do not guarantee growth for NSF and NIH.

But Jack Gibbons, Clinton's science adviser, said that, although there had been widespread concern that balancing the budget might require cuts of a quarter or one-third in research, "we've been working to make sure that does not come to pass".

Under the new projections, non-defence R&D would grow by 2 per cent between 1998 and 2002. But Gibbons stressed that research agencies might choose to fight to win larger increases than that.

Al Teich, head of science policy research at the American Association for the Advancement of Science, says that a combination of the revised projections and lower inflation — down to 2.2 per cent — is likely to ease the pressure on science funding.

In 1998, the administration proposes more funds for civilian technology programmes, including the Advanced Tech-

nology Program (up 22 per cent to \$275 million) and the dual-use programme at the Department of Defense (up 24 per cent to \$225 million). But these increases are less aggressive than those sought for technology programmes in previous Clinton budgets.

George Brown (Democrat, California), the senior Democrat on the science committee in the House of Representatives, said that Clinton's budget "falls far short of what needs to be done". Brown has produced his own "investment budget" which funds more science and education by cutting military spending. But he welcomed "the improved long-term outlook for agencies such as NASA" in the budget projections for the years beyond 1998.

Republican reaction to Clinton's budget proposal was muted, with congressional leaders leaving Washington the day after the budget for a weekend retreat to consider their strategy.

Science lobbyists detect growing signs that the Republican leadership will back funding increases for research. For instance, Senator Trent Lott (Republican, Mississippi), Senate majority leader, has called for more Department of Defense support for universities. In a letter to Ted Stevens (Republican, Alaska), the chair of the Senate appropriations committee, Lott praised the universities' role in national defence and attacked "the eroding commitment to research" by the department.

Colin Macilwain

For full budget details, see pages 568 and 569

Japanese increases include support for CERN's Large Hadron Collider

[TOKYO] Japan has increased its financial support for the Large Hadron Collider being built at the European Laboratory for Particle Physics (CERN) in Geneva, Switzerland, in a supplementary budget aimed at reinvigorating the flagging Japanese economy.

The Ministry of Education, Science, Sports and Culture will receive ¥3.85 billion (US\$31 million) for collider construction in a supplementary budget for fiscal year 1996, which ends on 31 March, passed by Japan's parliament, the Diet, at the end of last month. This comes on top of ¥5 billion given to the project in a supplementary budget in 1995.

The past few years have seen a string of supplementary

budgets, intended mainly to help pull Japan out of recession. But each has also contained substantial outlays for science and technology.

The present supplementary budget, at ¥2,666 billion, is small compared to its predecessors, some of which exceeded ¥10,000 billion. But, as anticipated after the Liberal Democratic party regained power in October (see *Nature* 383, 750; 1996), the outlays for science and technology are comparatively large, at ¥155 billion. This will add 5 to 10 per cent to the rapidly increasing budgets of science-related ministries and agencies.

Key beneficiaries will be Japan's universities, including the national research institutes used

by several universities, such as the high-energy physics laboratory at Tsukuba science city and the Institute of Space and Astronautical Science.

Japan's contribution to the international space station goes up by ¥7 billion, and there is ¥5.6 billion for SPring-8, the world's most powerful synchrotron. Another ¥3 billion will be spent on Mirai, the giant oceanographic vessel made by converting the ill-fated nuclear-powered ship *Mutsu* to diesel-electric power.

The Ministry of International Trade and Industry (MITI) will use its supplementary budget of ¥33.2 billion to improve facilities at the research institutes of its Agency of Industrial Science and

Technology: computer networks.

Another ¥2 billion will be spent on a system to monitor geological faults in the region affected by the Kobe earthquake. The Science and Technology Agency receives a similar amount to set up stations to monitor microearthquakes.

Chris Llewellyn Smith, the director general of CERN, said that he was "delighted" by the LHC decision. "The generous contributions of Japan, with those of other non-member states — particularly the United States — will enable the LHC to be built in a single stage and represent an important step forward in interregional scientific collaboration," he said.

Richard Nathan & David Swinbanks

Heat rises over UCSD 'misconduct' charge

[SAN DIEGO] In an unusual conflict between academic officials and federal authorities, the US Department of Justice has taken over a federal lawsuit against a researcher at the University of California at San Diego (UCSD) concerning allegedly fraudulent grant applications at the same time as an academic panel at the university has suggested that the scientist should be cleared of misconduct charges.

The case stems from three-year-old allegations that Maurizio Zanetti, a former Italian physician turned immunologist, fabricated research results and violated university policies covering the use of animals in research, human experiments and biosafety standards.

The allegations were made by Paolo del Guercio, another Italian former physician who held a junior position in Zanetti's laboratory, which specialized in AIDS research. Del Guercio lost his university post after making the allegations. Zanetti denies any scientific misconduct.

Zanetti, a tenured professor, was a close colleague of the late Jonas Salk, and carried out research with the Immune Response Corporation of Carlsbad, California, which

Salk set up to try to devise a therapeutic vaccine for AIDS.

Del Guercio claims in a 'whistleblower lawsuit' filed in 1995 that Zanetti made false claims to win between \$1.8 million and \$4.8 million in grants from the National Institutes of Health. Del Guercio previously carried out research for the French national medical research agency, INSERM, in Paris, as well as for the Wellcome Foundation in Great Britain.

In 1995, an *ad hoc* committee of UCSD scientists, acting in response to Del Guercio's complaints, found evidence that Zanetti had fabricated data in at least two published articles. The committee suggested that Zanetti might have suppressed results of an HIV experiment so that he could obtain an NIH grant.

But two months ago, a further faculty panel set up to consider whether Zanetti should be punished decided that he had not fabricated experiments or suppressed results. This second panel did, however, conclude that Zanetti had violated the university's guidelines on experiments, and recommended that he be given a strongly worded letter of reprimand.

The panel's recommendations were sent to Robert C. Dynes, the university's vice-chancellor, to decide on what action should be taken. A spokeswoman for the university says that Dynes will make "an informed" decision with the benefit of "substantial input" from the panel.

Meanwhile, however, a further dimension has been added by the contents of Del Guercio's lawsuit, which was unsealed in the US District Court in San Diego. This names Zanetti (as a principal investigator), the university and a junior researcher in the laboratory, Rosario Billetta, as having made false statements to the NIH to secure grants.

Most of the allegations in the lawsuit repeat those made to the university. The grants were for research on AIDS, a malaria vaccine and a vaccine for experimental allergic encephalitis. Billetta, another Italian researcher who recently left UCSD, has denied any impropriety.

Last month, the US Attorney's Office in San Diego revealed that it was intervening in the lawsuit, using its authority in such cases to take over the prosecution, and suggesting that the government may believe that some of the allegations are true. If the court decides that NIH grants have been improperly obtained, federal officials will seek to recover them from the university—with Del Guercio sharing in the recovered funds.

Neither the federal authorities nor UCSD officials are prepared to discuss the issue. But, in a highly unusual move, Zanetti has released all documents and transcripts of the two-week hearing before the second university panel.

These reveal Zanetti's attorney, Milton J. Silverman, who specializes in high-profile criminal defence work, conducting a harsh attack on the characters of his accuser and UCSD officials. Melvin W. Beal, an attorney for the university, said during the panel's hearing that Zanetti's defence "indicates an attitude of monumental arrogance and disdain for the principles that underlie the academic enterprise".

Beal also said that "whether dealing with the canons of intellectual honesty, the safety requirements relating to biological research, human subjects or animal use, Dr Zanetti shows himself to be either utterly without recognition of the importance of these principles or voluntarily unwilling to adhere to them".

But Silverman responded that there was "not a fraudulent bone" in Zanetti's body, describing him as "a good scientist and a decent man". Zanetti "is not perfect, he has made mistakes", said Silverman, but he "is completely innocent" of the scientific misconduct charges.

Rex Dalton

Rifts found as Antarctic ice breaks apart

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

[LONDON] Researchers from Greenpeace, the environmentalist group, last week came across significant rifts in the southern part of the Larsen ice shelf in the Antarctic peninsula — an area known as Larsen B (see above). Scientists from the British Antarctic Survey predicted last year that Larsen B was beginning to rift and that eventually the neighbouring sector, Larsen C, "may well behave in a similar way" (see *Nature* 378, 328; 1996).

The researchers found that five Antarctic ice shelves out of nine they had studied have disintegrated in the past 50 years.

Meteorological records over the same period show that the disintegration coincided with a 2.5° C rise in atmospheric temperatures.

Accelerated disintegration is a common feature of retreating ice shelves in the Antarctic. The final stage of disintegration of Larsen A, the 4,200 square-kilometre northern part of the Larsen ice shelf, was marked by sudden collapse in January 1995. Some 1,300 square kilometres of ice were lost in 50 days when the ice shelf broke into thousands of small icebergs, producing a plume 200 km into the Weddell Sea (see *Nature* 374, 108; 1995).

UK universities urged to join forces to boost research

[LONDON] Closer collaboration between British universities, aimed at creating world-class research centres and widening access to the latest research equipment, has been suggested by the government in a paper from the Department for Education.

But government officials appear to be backing away from the idea, which some were actively promoting last summer, that leading research universities should be identified as belonging to a 'superleague', with privileged access to research funds.

The government's suggestions come in a paper submitted by the department to a national commission that is looking into the future size and structure of higher education. The review is being chaired by Sir Ron Dearing, a former civil servant.

Both the Conservative and the Labour parties have indicated that, whichever wins the general election due to be held within the next three months, the next government is likely to accept many of Dearing's proposals.

The main focus of the inquiry has been to review the financing of universities, and in particular how many students universities should accept, and how much students should pay towards their education. But its remit includes all aspects of the way that higher education institutions operate.

On the question of research, the department's main suggestion is that serious consideration should be given to increasing the concentration of research funds and to promoting stronger links between universities and local communities.

According to the department's paper: "Any method of allocating public funding for research should allow for fruitful collaboration and the emergence of new groupings of researchers. The rising cost of support infrastructure gives a further impetus to this development."

In parallel with this, the department points out that the mounting costs of investing in research infrastructure argue in favour of the concentration of resources and against "significant turnover in the departments receiving research funding".

Such a move would have widespread support among the research councils. John Krebs, for example, chief executive of the Natural Environment Research Council, writes in a foreword to his council's annual report, published last week, that there is a "common perception" that "the limited resources available for supporting research are spread too thinly and that greater focus... is both desirable and inevitable".

According to the government paper, one implication of a strategy of concentration

could be "a higher threshold for research funding". Last summer, officials in the department used similar arguments to back suggestions for the creation of a 'superleague', to which a limited number of universities — some suggest a dozen — would be admitted (see *Nature* 382, 381; 1996).

The idea has considerable (though not unanimous) support from institutions that are confident they would be included in this league, and which feel that their research efforts are in danger of being undermined by government efforts to spread scarce research funds too broadly.

Last month, for example, the vice-chancellors of the Universities of Cambridge, Warwick and Edinburgh, as well as University College London, wrote an article in *The Observer* newspaper suggesting such a move. They say this is necessary to counteract a general decline in standards in some academic subjects, in particular mathematics.

An earlier government position paper appears to endorse this view, saying that "neither teaching nor research is best served by allowing all universities and colleges to pursue necessarily limited research funds".

But the official submission to the Dearing inquiry highlights some of the drawbacks of such a scheme. It points out that excessive barriers between those designated as 'research universities' and other institutions "could stifle competition, prevent the development of emerging departments and discourage research projects which, though not necessarily of world class, are important to the regional or local economy".

Elsewhere, the paper says that, if not all higher education institutions were allowed to compete for research funds — as they do at present — "it would have to be decided how eligibility for research funds should be determined, without entrenching a new binary line and fossilizing research capability".

The government's support for greater collaboration between universities to create joint centres of excellence and share expensive equipment is likely to receive considerable backing from universities themselves, which are increasingly concerned with making cost-effective use of scarce resources. "This is very much what we have been suggesting," says Michael Powell, secretary of the research committee of the Committee of Vice-Chancellors and Principals.

The apparent decision not actively to endorse a 'superleague' is also likely to win widespread endorsement, particularly from institutions that fear such a move might exclude them from applying for research support from public funds.

David Dickson

Congress turns the microscope on science spending

[WASHINGTON] Increased monitoring of federal science agencies, aimed at ensuring that taxpayers get value for money, was revealed last week to be high on the agenda of the new chairman of the House of Representatives Science Committee for the committee's activities over the next two years.

James Sensenbrenner, the plain-speaking Wisconsin Republican who takes over from the retired Robert Walker of Pennsylvania, invited scientists to help decide how the value of research and development spending should be measured. He said he is "strongly opposed to Congress micromanaging federal agencies" but wants to achieve a consensus on broad programme goals for scientific spending. "If we're successful," he said, many



Sensenbrenner: seeking consensus.

of the scientific controversies that plagued the last Congress "will die away".

Unlike his predecessor, Sensenbrenner does not foresee a single, cabinet-level Department of Science. But he does think that a unified federal science budget would be a "step in the right direction", as it would avoid pitting research spending against popular programmes such as housing and veterans' benefits, which are funded out of the same authorization bills.

The new chairman said he would like his committee to "look beyond the narrow issue of direct federal funding and management of science". He plans to hold committee hearings on corporate research and development to establish how government agencies can best use their investments to secure more private spending.

The new chairman, who formerly headed the committee's subcommittee on space, will start by focusing on subjects familiar to him. A priority is passing space legislation that went through the House last year but stalled in the Senate.

Sensenbrenner's first hearing, to be held this week, will be on Russian participation in the international space station project. He will travel to Russia and Europe next week with the new chair of the space subcommittee, Dana Rohrabacher (Republican, California), to assess the situation.

Other new Republican subcommittee chairs include Ken Calvert of California, who takes Rohrabacher's place as head of the subcommittee on energy and the environment; Constance Morella of Maryland, who will remain the chair of the technology subcommittee; and Steven Schiff of New Mexico, who stays as head of the committee's basic research subcommittee.

Tony Reichhardt

Congress vows to contest 'unacceptable' NIH plans

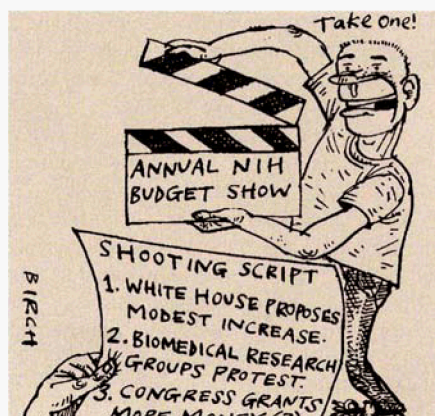
[WASHINGTON] Biomedical lobbyists are disappointed that President Clinton's budget proposals for 1998 would increase funding for the US National Institutes of Health (NIH) by 2.6 per cent, only slightly higher than the anticipated level of inflation.

If approved by Congress at this level, the result would be to raise NIH funding by \$337 million, to \$13.1 billion, with 80 per cent of the new money going to research project grants. The increase is substantially less than the 6.9 per cent and 5.7 per cent increases that Congress granted in 1997 and 1996 respectively. It is also slightly less than the 3.1 per cent 'biomedical inflation' rate in 1998 as calculated by the agency.

Nevertheless such an increase would enable NIH research project grants to reach a historic high, with 26,679 grants, a 3.6 per cent increase over 1997.

The budget includes \$90 million for continued construction of a \$310-million, 250-bed Clinical Research Center at the NIH in Bethesda, Maryland. It also boosts funding for the National Institute on Drug Abuse by 9 per cent — proportionately far more than other institutes — to \$358 million. The new money will be targeted at developing medication for cocaine addiction.

Donna Shalala, Secretary of Health and Human Services, said last week of the 2.6 per cent increase: "Is it enough? The answer is 'no.'" But she insisted that biomedical research is a "priority" and said the modest



increase is necessary because many other programmes also depend on the huge department's roughly \$35-billion budget for non-mandatory expenditures.

NIH's director, Harold Varmus, said the increase would allow the agency to "maintain [its] momentum". "It's a budget I can defend" to Congress, he said. Varmus said that healthy increases over the past two years, combined with a lower than expected rate of biomedical inflation, mitigate the impact of the modest increase. Inflationary increases for continuing grants will be held to 2 per cent, and research management and support will be kept at this year's level.

The overall increase drew loud complaints from biomedical research advocates. John Suttie, president of the Federation of Societies for Experimental Biology and a

biochemist at the University of Wisconsin, Madison, said: "What the NIH budget needs is continued inflationary adjustment, and then some meaningful increase."

Congressman John Porter (Republican, Illinois), chairman of the House of Representatives appropriations subcommittee that funds NIH, called the proposed 2.6 per cent overall increase "unacceptable". He said that recent statements by Clinton strongly supporting biomedical research were inconsistent with his budget. "If the president is going to talk about the importance of research, his budget ought to reflect that as a priority."

In his State of the Union speech last week, Clinton said: "We must reinforce our commitment to medical science." He called NIH "the most powerful discovery engine for an AIDS vaccine" and boasted that under his administration AIDS funding at the agency has increased "dramatically". The budget calls for \$1.54 billion for AIDS research in 1998, a 2.6 per cent increase over 1997, and a 44 per cent increase over 1993.

Within this, the budget for vaccine research will increase by 13 per cent, to \$148 million. "We feel there's greater potential now for thinking constructively about vaccines," Varmus said. The president's budget channels all AIDS money through the Office of AIDS Research, a policy that congressional Republicans oppose.

The Centers for Disease Control and Prevention, based in Atlanta, would receive \$2.3 billion, a 0.6 per cent increase. A programme dealing with emerging infectious diseases would receive a \$15-million increase, aimed in part at epidemiological and laboratory research. **Meredith Wadman**

Despite cutbacks, NASA pulls out of a threatened long-term nosedive

[WASHINGTON] Although the US National Aeronautics and Space Administration (NASA) is not scheduled for an increase in the proposed 1998 federal budget, it has been spared the steep decline in its long-term funding that was threatened last year. A relieved NASA administrator, Daniel Goldin, said last week: "I hoped we could do better" than last year's projection. "I am happy to report we have."

Instead of seeing its funding drop below \$12 billion in the next decade, as the White House had forecast a year ago, NASA is asking for \$13.5 billion for 1998, a total which is only slightly less than this year's final appropriation of \$13.7 billion. The agency now anticipates putting in a request for \$13.4 billion for 1999,

and \$13.2 billion for fiscal years 2000 to 2002.

The agency is seeking \$2 billion for space science — including astronomy and planetary exploration — in 1998, about the same as it received this year. The request for life sciences and microgravity sciences falls by 12 per cent, to \$214 million. In contrast, Earth science, including the agency's Earth Observing System satellites, increases by 4 per cent to \$1.4 billion.

The improved long-term budget picture will allow NASA to proceed with several new initiatives, including the Space Infrared Telescope Facility and an average boost of \$50 million a year for spacecraft technology development, which will push the schedule for a Mars sample return

mission forward to 2005.

Additional technology spending in the next few years will also allow spacecraft missions to Pluto and Jupiter's moon Europa. Wesley Huntress, head of science at NASA, says that one of the two will be selected for funding from the beginning of 2000.

The agency also kicks off its new 'Origins' programme in 1998, which includes a wide range of projects from the infrared telescope to basic research in extraterrestrial biology. These and other new science initiatives will be financed largely through savings in the space shuttle programme and agency overheads, spending on both of which will continue to decline. Funding for the space station

remains capped at about \$2.1 billion a year.

The higher five-year budget projections have killed interest in a 'space summit' meeting between the White House and Congress to decide NASA's long-term spending outlook (see *Nature* 384, 610; 1996). Goldin credits Senator Barbara Mikulski (Democrat, Maryland), who called for the summit meeting last year, with forcing the Clinton administration to re-evaluate its projections for the agency.

The new chairman of the House of Representatives Science committee, James Sensenbrenner (Republican, Wisconsin) said last week he thought the summit would be a "waste of time" until questions about the space station are resolved. **Tony Reichhardt**

NSF greets 'a good budget'

[WASHINGTON] The National Science Foundation (NSF), the main source of funds for non-medical research in US universities, would receive an increase of 3 per cent under President Clinton's proposals for the next financial year, increasing its budget to just under \$3.4 billion.

Science lobbying groups, which have set their sights on securing an increase of at least twice that amount, were quick to attack the proposal as inadequate. But Neal Lane, director of the foundation, said that it was "a good budget for the NSF, given the constraints" on overall government spending.

Under the proposal, the NSF will provide \$9 million for design work on the Millimetre Array, a proposed millimetre-wavelength telescope, and \$25 million for the Polar Cap Observatory near the North Pole. The proposals also include \$25 million towards renovation of NSF's ageing research stations at the South Pole, and \$10 million towards an initiative, personally backed by Clinton, to develop a 'second generation' Internet.

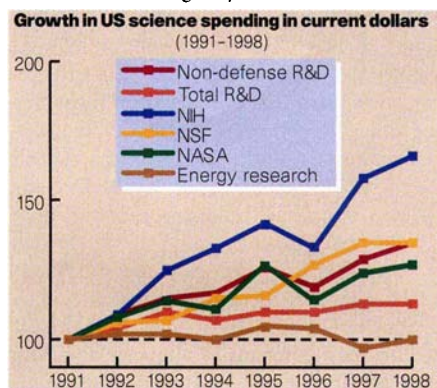
Lane declined to specify what NSF would have to give up to fund these new initiatives, merely saying that "we'll stop doing good things in order to do better things". But he promised to "discuss with the community where these cuts will take place".

The increase in the foundation's budget, which is basically sufficient to cover inflation, now running at 2 per cent in the

United States, will be spread evenly across the NSF's three main areas of activity — research grants, research facilities and education. Of the agency's seven directorates, social sciences and computing fare best, with increases of 6.5 per cent and 7.6 per cent respectively. In contrast, geosciences lags with a proposed increase of just 1.5 per cent.

Cornelius Pings, president of the Association of American Universities, which earlier in the year joined with other science lobbying groups in calling for a 7.1 per cent increase for the NSF (see *Nature* 385, 103; 1997), says that he plans to continue to push the Congress for a larger increase.

Pings says he is "reasonably optimistic that we could end up with something more aggressive than what the president has put forward" for the agency. **Colin Macilwain**



Green light sought for laser fusion and LHC

[WASHINGTON] Full funding for two major physics projects is included in the US energy department's budget proposal — the National Ignition Facility at Lawrence Livermore National Laboratory in California, and the US contribution to the Large Hadron Collider at CERN in Switzerland.

The administration took the unusual step of asking for the entire \$1 billion needed to complete the National Ignition Facility — which will test inertial confinement fusion — to be given approval by Congress in the 1998 financial year. At the same time, it is seeking \$450 million for the collider over eight years under an 'advance appropriation' by which Congress would agree in principle to support the project to completion.

The collider request follows the conclusion last week of a framework agreement on the US contribution to the project. The deal was initialled by Martha Krebs, head of the energy department's Office of Energy Research, and Christopher Llewellyn Smith, director general of CERN, the European laboratory for particle physics in Geneva (see *Nature* 385, 103; 1997).

The department plans to increase its spending on research on detectors for the

collider from \$15 million to \$35 million in 1998. The total budget for high-energy physics will remain virtually level at \$675 million. But money for the collider and other programmes will be released by this year's completion of the B-Factory at Stanford Linear Accelerator Center in California.

Nuclear physics will receive a budget increase of 5 per cent, to \$332 million, and basic energy sciences an increase of 4 per cent, to \$672 million. The latter includes \$23 million for the design of a proposed spallation neutron source at Oak Ridge, Tennessee.

In contrast, the administration is proposing to reduce its support for fusion science by 4 per cent, to \$225 million. But this total does include \$13 million towards the construction of a compact tokamak at Princeton Plasma Physics Laboratory in New Jersey.

Elsewhere, the administration found an extra \$200 million to boost research into solar and other renewable energy sources, as well as \$100 million for nuclear energy. Environmentalists welcomed the former, but criticized the latter as "wasteful and unnecessary corporate welfare". **C.M.**

EPA seeks money for chemical clean-up

The US Environmental Protection Agency is asking for \$7.6 billion next year, a 12 per cent increase over its 1997 budget. Much of that will be used to accelerate the clean-up of polluted chemical disposal sites. Science and technology spending will also grow significantly by 12 per cent, to \$614 million.

A considerable portion of the new science spending will go to strengthen research and monitoring of air pollutants, including ozone and fine particles. The agency plans to increase its spending slightly on grants to outside scientists to \$100 million.

Biologists buck trend in Geological Survey

The newly formed Biological Research Division, the successor to the US National Biological Service, is one of the few areas within the US Geological Survey to receive increased funding. The request for the division goes up slightly to \$145 million; overall the survey is seeking \$745 million for 1998, \$6.5 million more than it received this year. Support for the survey's geological, mapping and water resources divisions would remain essentially flat.

Basic research wins boost in defence

The budget proposal increases basic research funded by the Department of Defense — most of it conducted in universities — by 5 per cent, to \$1.19 billion. It also boosts funding for the Dual Use Application Programme, which develops technologies for both military and civilian use, from \$181 million to \$225 million. But, controversially, it cuts almost \$800 million from the \$35 billion that the department spends on developing equipment.

Technology programme cuts back ambitions

The Advanced Technology Programme operated by the National Institute of Standards and Technology would be raised from \$225 million to \$275 million, and the Manufacturing Extension Partnership from \$95 million to \$123 million, under the budget proposal. But the administration has shelved its previous ambitious plans to expand the ATP to a \$1 billion programme.

The budget requests flat funding for the institute's laboratories, and drops plans to replace lab buildings at Gaithersburg, Maryland and Boulder, Colorado. The National Oceanic and Atmospheric Administration, the other science agency in the Department of Commerce, is seeking an increase of 4 per cent, to \$2 billion.

'Medicinal plants threatened by over-use'

[LONDON] Urgent measures are needed to cultivate endangered medicinal plants, some of which are potentially threatened with extinction in the wild as a result of the growing popularity of herbal medicines, according to a report by a London-based management consultancy.

The report identifies 18 species "in urgent need of conservation", and forecasts a doubling of the quantity used of 12 of the species within the next decade unless conservation measures are taken. It calls for measures to cultivate threatened plants, particularly in India and China, two important markets for medicinal plants, where they are gathered largely from the wild.

Sales of medicinal plants have grown by nearly 25 per cent in India over the past ten years, the highest rate of growth in the world. Germany is Europe's largest consumer of such plants, and spends £1.4 billion (US\$2.2 billion) annually. France is second (£116 million) and the United Kingdom third (£88 million). Products used in cardiovascular and muscular disorders are popular in

Europe, while primary health-care products are more popular in developing countries.

"It is ironic that the growth and popularity of herbal products are now threatening to destroy the source of these popular remedies," says Gopi Warriar, co-author of the report and chairman of McAlpine, Thorpe and Warriar, a management consultancy that specializes in analysing the worldwide market for herbal products.

Warriar says commercial organizations need to take a closer interest in conservation. He calls for the formation of an association of "powerful herbal companies" independent of governments.

Although largely welcomed by scientists and companies involved in growing, distributing and manufacturing herbal products, the report has caused some controversy — and not merely because of its price of £890. Some, such as Ghilleen Prance, director of the Royal Botanic Gardens at Kew in London, point out that conservation measures are under way through the United Nations Convention on Biological Diversity, as well as a recent UN Food and Agriculture Organization declaration on the conservation of plant genetic resources.

Also at issue is the inclusion of *Ginkgo iloba*, used in the treatment of cardiovascular illnesses, on the list of "species in urgent need of conservation". "Whoever says this plant is threatened is completely off base," says Marlin Huffman, a senior partner in Plantation Medicinals of Felda, Florida, one of the two largest US herbal companies, with an annual turnover of \$20 million.

Ginkgo originated in eastern China. It was once common near Buddhist temples, but is now virtually extinct in its native habitat, says David Chamberlain, curator of the herbarium at the Royal Botanic Garden in

Edinburgh. But he points out that the plant is still commonly found in Europe and the United States, where it is widely cultivated.

The report's authors nevertheless argue that the supply of *Ginkgo* may not be able to keep pace with demand, which is increasing at 26 per cent a year, one of the highest for any medicinal plant. Two thousand tonnes are used each year, mostly in Germany. It is second only to *Panax ginseng*, whose present demand at 8,000 tonnes a year is projected to grow at 16 per cent a year.

But the report's suggestion that producers should liaise with local scientists to gather data on threatened species has been welcomed by scientists such as Vernon Heywood, president of the International Council on Medicinal and Aromatic Plants.

Compiled over 18 months, the report is based on data on the status of plant conservation from the World Conservation Monitoring Centre in Cambridge, England, combined with the company's own forecasts of the supply and demand of medicinal plants.

The report identifies another species, *Saussurea costus*, which is rare in the wild, but demand for which in gynaecology and as an immunity-enhancer and ingredient in cosmetics is likely to increase. *Harpagophytum procumbens*, a plant popularly used for muscular disorders, is also under threat, but few data exist on its conservation status. This plant, too, is gathered straight from the wild, with little or no domestic cultivation. More than half of the 600–800 tonnes sold annually go to Germany. Demand is expected to quadruple in the next decade.

The report says that urgent conservation measures are needed in China and India, where up to 80 per cent of people rely on herbal formulations for primary health-care. Two plant species in India, *Adhatoda beddomei*, used as a diuretic, and *Inula racemosa*, used for stomach disorders, are particularly vulnerable.

But Huffman believes that conservation measures will be difficult to achieve in practice so long as the returns made in developing countries from selling wild plants remain relatively high. "If a man can make a few dollars a day, instead of a few cents, then he'll go right ahead and dig the plant."

Huffman says the report, which also calls for western companies to collaborate with local companies, particularly in developing countries, will be useful to producers. The idea of the forum has been welcomed by another US herbal plant company, Trout Lake Farm in Trout Lake, Washington state. But Brian Rust, product procurement director, says cultivating plants outside their natural habitat is not easy, and can be costly because soil and environmental conditions cannot be exactly duplicated. **Ehsan Masood**

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Ginkgo: demand for the plant, almost extinct in the wild, is growing at 26 per cent a year.

Breach of embryo ban leads to resignation

[WASHINGTON] A leading reproductive biologist has resigned from his post as director of the Institute for Molecular and Human Genetics at Georgetown University in Washington DC after admitting that he breached the Jesuit university's ban on embryo research.

Mark Hughes' resignation, tendered in a letter to the university at the end of last month, follows the termination last October by the National Institutes of Health (NIH) of a contract under which Hughes had been carrying out research at the National Center for Human Genome Research (see *Nature* 385, 190; 1997).

The NIH cited Hughes' breach of a government ban on federal funding for embryo research. Hughes used NIH-funded

fellows and equipment to carry out pre-implantation diagnosis on test-tube embryos at Suburban Hospital in Bethesda, Maryland. He has since admitted that he also conducted some of this work on the Georgetown campus.

"Recent events make clear that I cannot continue my pre-implantation genetic diagnosis work while serving as a Georgetown faculty member," Hughes wrote. He said he would resign immediately rather than give up work he considered to be of "crucial importance".

Hughes is involved in identifying disease genes in embryos. These embryos are subsequently discarded before implantation, allowing apparently healthy embryos to be implanted instead. **Meredith Wadman**

France takes gentler approach to reforms

Declan Butler

By consulting widely on changes in the university sector, France has so far avoided the large-scale protests that met previous attempts.

The familiar sight of students and lecturers protesting against university reform on the streets of Paris may be heading for the history books. Last week, for the first time in decades, the French government was able to announce a programme of reform that met with wide consensus rather than resistance.

One possible explanation is that the government has jettisoned its old approach of imposing change from above in favour of broad consultation with all the parties. Another is that the conservative coalition government seems to have abandoned its traditional aim of replacing France's Napoleonic university system, with its cherished principle of open access to all suitably qualified students, with an Anglo-Saxon model of competitive research at universities with limited student access.

The reforms are the first to emerge from a national consultation process launched by education minister François Bayrou, and aimed at producing a government-endorsed white paper (policy document) on the university system by the end of the year.

The most significant reform is intended to reduce failure rates in the first two years of study leading to a university diploma of general studies. Only 60 per cent of students pass the course, and only one in four of these within the two years, the rest having to resit.

In future, students will receive a more general education in the first semester, enrolling in one of eight broad courses instead of specializing from the outset.

Ambitious construction plans

Other reforms include streamlining the three-semester academic year into autumn and spring semesters, as in US universities, and an increase in tutorials. There is also a promise to extend the ambitious Université 2000 construction plan — launched in 1991 by Lionel Jospin, the then education minister — under which France has already spent FFr32 billion (US\$5.8 billion) over the past five years to build more than 1.5 million square metres of university space.

The outcome of the first phase of the consultation process has been broadly welcomed by university presidents and students' and teachers' groups, although all are reserving final judgement until specific

details are made available about how measures will be implemented and funded. "Bayrou has in the past made promises that were not kept," warns Laurent Bresset, secretary general of the main body representing teachers, the SGEN-CFDT.

This first round of reforms is hardly revolutionary. What is new is the commitment to this and further reform that Bayrou seems to have won from students and staff. In the past, resistance to change from powerful lobbies of students and *professeurs* has impeded reform efforts. "The changes seem so obvious that you might ask yourself why we didn't do this 15 years ago," says one observer.

Bayrou — a former schoolteacher who is seen as a future presidential candidate — appears to have succeeded by consulting widely with all the parties over the past 18 months and hammering out points of consensus. 'La méthode Bayrou', as it is widely called, is the only way to make progress, acknowledges Hélène Lamicq, president of the University of Paris 12-Val-de-Marne.

Yet another reform imposed from above would have been "doomed to fail, no matter what its quality", she says. The government's shift to seeking consensus marks a "real change in a centralized state like France", she says. Bresset says: "It's the first time there have been few criticisms."

Another major change is that the conservatives seem to have abandoned their long-standing goal of creating autonomous universities able to set their own course structures, curricula and fees. At present all universities follow national norms. "This is a historic shift in the policies of the right-wing parties," says Bresset.

Similarly, Bayrou's affirmation of the Centre National de la Recherche Scientifique (CNRS) as a partner of the universities indicates that the government has abandoned plans to dismantle the agency and transfer its laboratories to the universities, says Maurice Gross, CNRS director of university relations.

The government now accepts the concept of a national science community in which both CNRS and the universities have their place, says Gross, providing an environment more conducive to cooperation. "There is a consensus that reform of the universities must take place without endangering organizations that work well," says one observer.

But the task of carrying through this

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Times past? Student protests have been avoided by opting for consensus, not confrontation.

reform remains enormous. Many universities are in a dire state of decay; classrooms are overcrowded, and one in ten buildings fall below minimum safety standards (see *Nature* 378, 651; 1995).

Are student fees the way ahead?

Lamicq says she hopes the improved relations between government and the universities will allow the "taboo" subject of student fees to be raised. Many still fear that such a discussion would open a Pandora's box, with universities pressing for fees commensurate with their facilities coming into increasing conflict with critics who argue that substantial fees would contradict the principle of open access.

At present, the state pays 80 per cent of the budget for the universities. But, with general pressure on public spending, the only way universities can boost their income is to increase student fees. Lamicq is convinced that students could be persuaded to accept increases, as long as they were incremental and across the board.

The dilapidated and overcrowded conditions at universities are in stark contrast to the alternative, highly competitive *grandes écoles*, where students enjoy excellent facilities, including well equipped libraries and laboratories.

In France, the cream of those leaving school undergo a special year of training — *la classe préparatoire* — for the entry competitions to the *grandes écoles*, while the rest go to university. Yet almost no one, even the students' unions, sees any contradiction between the general acceptance of the *grandes écoles* and widespread rejection of competitive entry to universities.

"Students accept all selection as long as it is based on merit," explains Bresset, who points out that access to university is automatic for students who succeed in winning a place by passing the school-leaving examination, the *baccalauréat*. Being excluded from a *grande école* is acceptable, says one observer, but being excluded totally from higher education is not. □

Canada calls in police to investigate data loss in HIV blood scandal

[MONTREAL] Canada's federal government has asked the Royal Canadian Mounted Police to investigate the destruction of key documents relating to the country's scandal over HIV-contaminated blood.

The Information Commissioner, John Grace, revealed last month that federal health department officials had destroyed documents following an enquiry under the Access to Information Act from a newspaper reporter.

The result has been the loss of transcripts and tapes of meetings of the Canadian Blood Committee, which oversaw the blood system between 1982 and 1989. During that period thousands of Canadians became infected with HIV and hepatitis C.

Grace's report said the committee had been under pressure from the Canadian Red Cross Society to avoid releasing documents that could be used in lawsuits filed by individuals who had been infected. Grace said that the Access to Information Act does not provide sanctions against those found to have improperly destroyed documents.

The commissioner's report has been forwarded to Mr Justice Horace Krever, who is heading a national inquiry into the blood

distribution system. Krever's report is due on 30 April. Victims' groups have demanded that criminal charges be brought against those who destroyed the evidence.

Russian science panel undertakes audit

[MOSCOW] Russia's state committee on science is to prepare an inventory of the country's 420 scientific organizations to find out their current activities, the scientific level of their research, whether they require state support and in what form.

As a result of this review, some institutions will be given state accreditation, bestowing certain privileges. But progress will depend on the adoption by the state Duma of a law on tax privileges for the accredited scientific organizations. This law is under discussion in committee, and will need to be approved by President Boris Yeltsin.

'Britain should repay £1bn bill for BSE'

[PARIS] The European Parliament's committee of inquiry into the handling of the bovine spongiform encephalopathy (BSE) epidemic last week withdrew its threat to table a motion censuring the European Commission (see *Nature* 385, 101; 1997). Such a move would have created a political

crisis if it had been backed by the full parliament. The committee rejected the proposal by 16 votes to three.

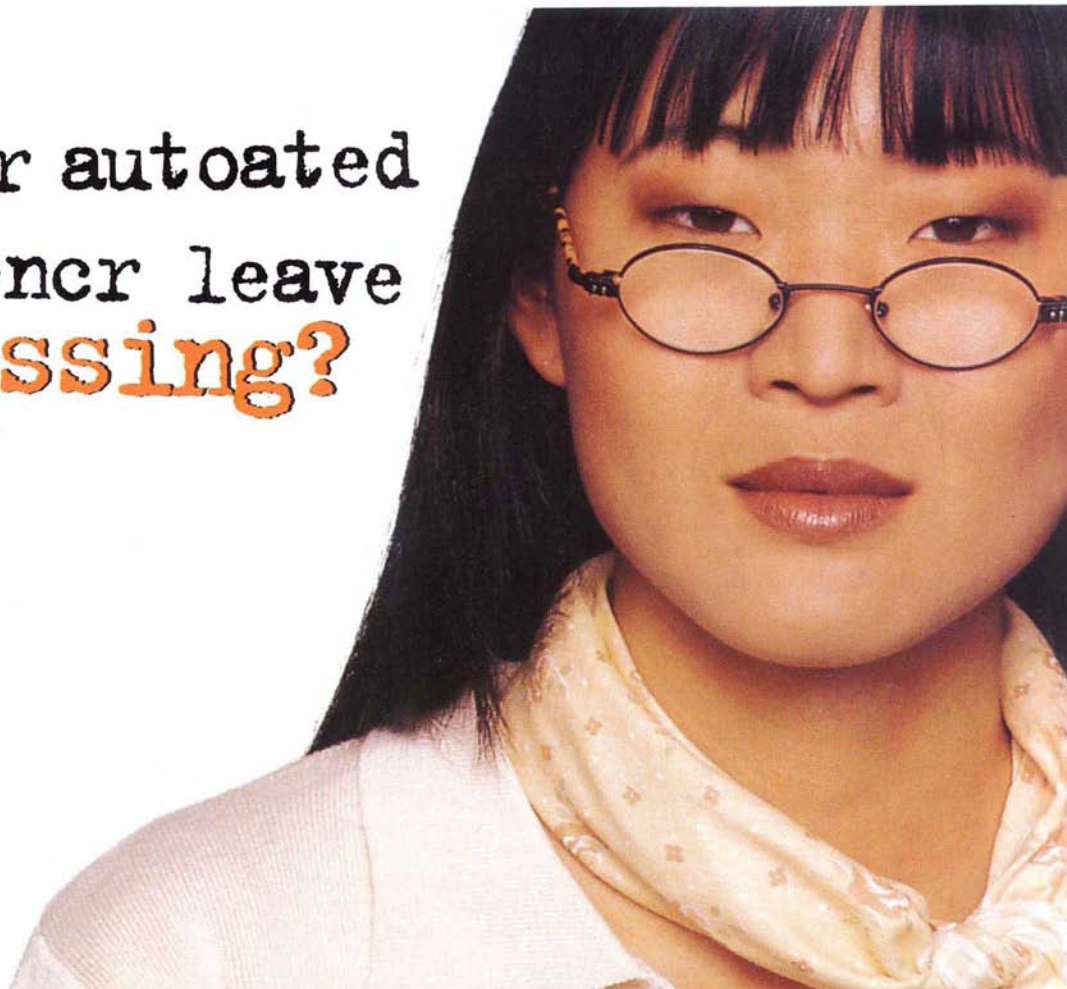
The committee's final report on the handling of the BSE crisis will be submitted to a plenary session of the parliament next week. The report accuses Britain and the commission of "misdemeanours and failures", but stops short of alleging a deliberate cover-up. The committee says that Britain should be forced to repay more than £1 billion (US\$1.6 billion) it has received from the European Union as subsidies to farmers since the start of the crisis.

India still hopes for Soviet reactors

[NEW DELHI] India has accepted an assurance that Russia will honour an eight-year-old commitment to supply two 1,000-MW light-water nuclear reactors, despite efforts by the United States to block the sale. The \$1.7-billion contract, under which the reactors were to be built as a 'turnkey' project, was signed in 1988. But the deal failed to go through following the suspension of talks after the collapse of the former Soviet Union.

Negotiations were revived three years ago after Russia agreed to build the reactors as a collaborative project at 70 per cent of the initial cost. Talks had been progressing

Does yur autoated
DNA seqencer leave
u **guessing?**



smoothly until the US government warned last week that the deal violated a 1992 agreement signed by the 'nuclear suppliers group', of which Russia is a member. But Russia has assured India that US objections will not affect the proposed sale.

Chernomyrdin agrees to space station plan

[WASHINGTON] US Vice-President Al Gore and the Russian Prime Minister, Viktor Chernomyrdin, announced last week they have worked out "in intricate detail" a plan under which Russia will meet its financial obligations to the international space station. The announcement followed increasing concern from the National Aeronautics and Space Administration (NASA) and its congressional supporters that problems with Russia's contribution to the station could mean an expensive bail-out by the United States.

Tokyo university post brings close contest

[TOKYO] Shigehiko Hasumi, vice-president and former dean of the general education faculty at Tokyo University, has won a closely fought election battle for the presidency of the university, beating Akinori Suzuki, another vice-president who is a biologist

from the faculty of agriculture. Suzuki won the first three ballots in which the university's 1,300 faculty members voted for various candidates. But in the final ballot Hasumi, an expert on French cinema, won after other candidates were eliminated.

Hasumi succeeds Hiroyuki Yoshikawa, an engineer who, like his predecessor, Akito Arima, has campaigned vigorously for reform of Japan's university research system. Some observers fear that the momentum for reform may gradually be lost with the change in leadership at Tokyo University, which takes effect on 1 April.

Researchers refuse to go gracefully in France

[PARIS] Elder statesmen in the French research community are up in arms following the government's decision to remove their right to continue working beyond the official retirement age. A new law on the employment of civil servants removes the right of scientists graded as 'exceptional' to retire at the age of 68, rather than 65.

The provision will affect about a hundred researchers at national research organizations, but will not apply to those working in universities. It will save the government around FF150 million (US\$27 million) between 1997 and 1999. Luc Montagnier, a member of the French group

that discovered the AIDS virus, is one of those affected.

Montagnier and several other scientists have formed a 'collective' defence group to fight the move, and defend "the general interest of French research". But a spokesperson for the science ministry says that, although the move would result in researchers being paid a pension instead of a salary, it does not prevent them from continuing to work.

Muses-B satellite aims for radio sources in space

[TOKYO] The Muses-B radioastronomy satellite (pictured at the launch pad), part of an international space observatory programme, was due to be launched by the Institute of Space and Astronautical Science in Japan on Tuesday (11 February). Muses-B will deploy an eight-metre diameter

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

telescope, which will operate in collaboration with more than 25 ground-based telescopes to observe radio sources in space, using very long baseline interferometry.

Does your automated DNA sequencer leave you guessing? If so, chances are it's primarily designed for high throughput sequencing. Why be uncertain of your sequencer's accuracy, when ALFexpress™ is providing researchers with the full genetic stories of their DNA.

ALFexpress: for more accurate readings

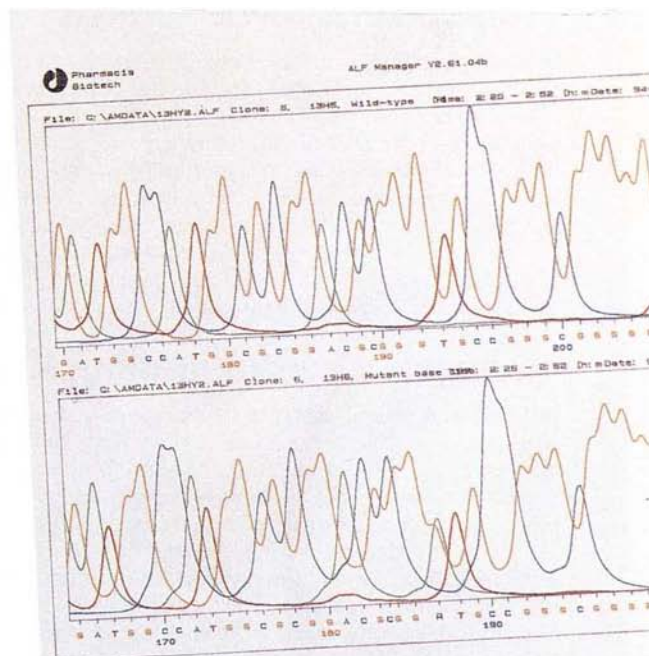
In the largest clinical study using automated DNA sequencing, the technology behind ALFexpress proved exceptional (see caption). That's one of many examples of ALFexpress offering unrivaled accuracy during automated confirmatory sequencing. Further, its readings are so accurate that ALFexpress can unambiguously identify heterozygous point mutations—as proven in many clinical research applications, such as analysis of tumor genes and high-resolution HLA typing. What's more, Pharmacia Biotech has dedicated software programs to support these applications.

For the full story, call us: 1 (800) 526-3593 from the U.S.; +81 (0)3 3492 6949 from Japan; or +46 (0)18 16 50 00 from Europe and the rest of the world. Or visit us on the Internet: <http://www.biotech.pharmacia.se>.



**Pharmacia
Biotech**

Uppsala, Sweden. (And the rest of the world)



The p53 gene from 316 breast cancer patients was sequenced using ALF automated sequencing technology. (Bergh J., Norberg, T., Sjögren, S., Lindgren A., Holmberg, L. "Complete Sequencing of the p53 Gene ..." *Nature Medicine* 1995; 10:1029-1034.)



Canadian Red Cross seeks to reform itself

Sir — Your News story, “Canada’s HIV blood inquiry turns poisonous” (*Nature* 384, 602; 1996), needs some clarification.

Your comments about the inquiry and the way it has been handled by Mr Justice Horace Krever in seeking to allocate blame for the infection may be true — it seems that way. The facts will come out on 30 April 1997 in his final report. I hope, however, that the inquiry will provide constructive information from which to set guidelines, not only for Canada but also for the rest of the world.

I am disturbed by the comments made by Dr Bert Aye, especially about the physicians’ roles at the blood centres and in the national office of the Canadian Red Cross. The transformation that is now taking place in the Red Cross is to a system comparable to the Canadian hospital model. The job descriptions and the roles of the physicians are in the process of being

clearly spelled out, and will reflect this hospital model.

Similarly, at the national office in Ottawa, the associate national director, medical and scientific affairs (in which position I am acting), has a major role in the area of medical development, scientific development, safety and education.

Several other key positions are also held by physicians. Two of the three regional general managers are physicians with expertise in transfusion medicine, and the third position is currently vacant. In addition, there are three regional medical officers who report to this office for medical and scientific purposes. These are physicians with expertise in transfusion medicine and they organize regionally the activities of patient services and they are part of the executive team at the regional office for manufacturing activities.

Furthermore, at the local centre level,

there is a management team formed by the centre manager, the quality assurance specialist and the medical officer. This management team deals with safety, medical and financial issues. Safety has definitely not been compromised, as safety, operational and medical issues are worked on by a team effort.

What has been happening during the past five months has been a move to develop and strengthen, under the influence of professionals, at national, regional and local levels, some of the weaker areas. These have included education, research in biotechnology, and forming alliances with agencies to strengthen our epidemiological surveillance.

It was not considered necessary to await the much delayed final report from Mr Justice Krever. Rather, we have moved forward in anticipation that our strategies will reflect or complement the recommendations to be made in that report.

Antonio Giulivi

*Canadian Red Cross Society,
1534 Blohm Drive,
Ottawa, Ontario, Canada K1G 4P9
e-mail: agiulivi@redcross.ca*

Freedom of choice

Sir — In the articles by Declan Butler and David Dickson in the “Science and risk: 1996” feature (*Nature* 385, 6–11; 1997), you comment on various aspects of risk evaluation in food safety in the European Union and the United Kingdom. There is a great deal of common ground among those considering difficulties in this area, and there is wide support for broader representation in the advisory process and for the involvement of views of groups other than those directly involved in the agricultural and agrochemical industries, food preparation and marketing. But the way in which this is done is critical, and, if Derek Burke is correctly quoted, I would be concerned about what seems to be a confusion between the roles of preparation of scientific assessments and the decision-making process.

The fact that consumers have “rejected” irradiated foods despite expert scientific advice that they are safe seems to me exactly what one would hope might happen in an informed and open society. The advice was presumably arrived at by a competent group with an adequate database and formed the basis of a decision that such foods were acceptable in health-care (and other) terms. That this advice was rejected means that, whatever the perceived hazard, it was not acceptable to a wide section of the community. This does not alter the

outcome of the scientific evaluation, which will change (or not) only when new, or different, data are available.

Modification of science-based advice by other inputs should come at a stage that is clearly distinct in the process of risk assessment. Advisory committees are advisory, and their advice (often good advice) can be ignored. The rejection of good advice is a phenomenon familiar to all medical practitioners (and many spouses) and a sign of healthy scepticism.

Colin Berry

*St Bartholomew’s and the Royal London
School of Medicine and Dentistry,
Royal London Hospital,
Whitechapel, London E1 1BB, UK*

Patronymy and conservation

Sir — Patronymy is an established practice in taxonomy, celebrating the patrons and pioneers of natural history exploration in species names. The description of a new bird at the end of last year (*Ibis* 138, 611–619; 1996) now heralds patronomy as a conservation tool.

The new species, the remarkable Chocó vireo (*Vireo masteri*), is found only in the threatened rainforest of the Colombian Andes. With this in mind, Paul Salaman and Gary Stiles, the discoverers of

the vireo, helped to found the pioneering BirdLife International Species Sponsorship Programme, and auctioned the immortality of the bird’s name. The honour went to Dr Bernard F. Master, on his donation of \$70,000 to establish a community nature reserve at Río Nambi, protecting a large portion of the bird’s remaining habitat (*Nature* 383, 661; 1996).

The implications of this initiative should be far-reaching for both conservation and taxonomy. The auction of the names of the few dozen charismatic vertebrate species discovered each year could net a tidy \$5 million annually. This figure could easily be exceeded if the names of the numerous but undescribed living (and fossil) invertebrates, plants and other organisms were auctioned for, say, \$1,000 each. Not only will this provide a much needed additional source of conservation funding, it should also stimulate the difficult and critical task of cataloguing our planet’s biodiversity. All credit goes to the vireo’s discoverers and patron, and to BirdLife International, for providing this opportunity. It is now up to taxonomists worldwide to make the most of it.

Thomas Brooks

*Department of Ecology and Evolutionary Biology,
569 Dabney Hall,
University of Tennessee,
Knoxville,
Tennessee 37996-1610, USA
e-mail: tbrooks@utkux.utcc.utk.edu*

Parity and chivalry in nuclear physics

Forty years ago, the world of physics was stunned by the discovery that nuclear beta-decay does not respect symmetry between left and right. But the credit for this conclusion has not been properly attributed.

Nicholas Kurti and Christine Sutton

On 15 January 1957, the physics department of Columbia University in New York held a press conference, and the following day headlines on the front page of *The New York Times* declared “Basic concept in physics reported upset in tests. Conservation of parity in nuclear theory challenged by scientists at Columbia and Princeton Institute.” It was unusual in those days for scientific results, however important, to be announced first at a press conference rather than at a scientific meeting or in a professional journal. The scientific paper appeared a month later on 15 February 1957 in *Physical Review*, and this fortieth anniversary of publication provides a good opportunity to look at those controversial events.

The great excitement was triggered by results from an experiment suggested the previous summer by theorists T. D. Lee and C. N. Yang to test the conservation of parity in weak interactions. Lee, from Columbia University, and Yang, from the Institute for Advanced Study in Princeton, had been considering one of the most puzzling effects of the time, the so-called ‘ θ - τ ’ paradox. θ and τ are two short-lived ‘strange’ subatomic particles, so called because although they are made readily in particle collisions through strong interactions, they decay only on the longer timescales typical of weak interactions. The additional puzzle was that they have exactly the same mass, but can decay to states of different ‘parity’.

Parity refers to a complete inversion in space, which has the effect, for example, of turning a right-handed corkscrew into a left-hand one. A system that has right-left symmetry will not change under the parity operation and is said to have positive parity, whereas a handed system, such as a corkscrew, does change and is said to have negative parity. The θ and τ decayed to sets of particles with opposite parity (two pions and three pions respectively), implying that they could not be the same particle, unless parity could somehow change in the weak decays. However, physicists believed that parity is conserved (stays the same) in all basic physical processes.

Lee–Yang hypothesis

In 1956, as more conventional explanations for the θ - τ puzzle were unsuccessful, Lee and Yang began to consider seriously the alternative of parity non-conservation and set out to discover whether there were other weak interactions in which parity was not conserved. After working through a great deal of

experimental evidence from nuclear beta-decay, they realized that no experiment had measured an effect that would change hands under parity. The key would be to measure something that changes, such as momentum, relative to something that does not, such as direction of spin.

In June 1956, Lee and Yang submitted a paper to *Physical Review* in which they suggested ways of testing for parity conservation¹. One idea was to measure the electron emission in β -decay from oriented nuclei to see if the intensity changed when the polarizing field was reversed. Any asymmetry would be proof of parity violation.

The NBS test

In the United States, the National Bureau of Standards (NBS) was one of only two places where nuclear orientation work, by Ambler, Hudson and Temmer, was being done. C. S. Wu, a colleague of Lee’s at Columbia University and an authority on β -decay, approached Ambler, who three years earlier had polarized γ -decaying cobalt-60 (^{60}Co) nuclei in his thesis work at the Clarendon Laboratory in Oxford. As a result it was decided late in July that the parity experiment would be carried out at NBS by Ambler, Hudson, Hayward (an experimental nuclear physicist), Hoppes (his research assistant), all from the NBS, and Wu.

Although the technique of γ -ray nuclear orientation experiments at millikelvin temperatures was well established by 1956, β -decay experiments presented special difficulties. Because of the strong absorption of β -rays, the scintillation detector had to be inside the cryostat and the light pulses transmitted to a photomultiplier at room temperature. For the same reason, the β -activity had to be concentrated in a 50- μm outer layer of the single-crystal cooling substance. To prevent this exposed layer from warming above the temperature of the bulk of the crystal, the thermal insulation had to be much better than in the γ -ray experiments. Thus it was not until December 1956 that the forward-backward asymmetry of the β -emission, on reversing the polarizing field, was first observed. The reason for describing these difficulties is that at the time many people thought that there was no more to the experiment than for Wu to turn up at NBS and receive from Ambler and Hudson a cerium magnesium nitrate crystal doped with ^{60}Co —and parity conservation tumbled.

Indeed, in his Rutherford memorial lecture in 1958, P. M. S. Blackett said “It took Wu and colleagues 48 hours to show

experimentally that the β -particles emitted from magnetically oriented ^{60}Co nuclei were emitted asymmetrically with regard to the direction of the magnetic field.” But he also pointed out that experimentalists were discouraged by the theorists from wasting their time on experiments that were bound to confirm parity conservation².

On 12 January 1957 the paper was ready to be sent to *Physical Review*. In those days it was usual to list authors in alphabetical order, unless one was the leader of the team or the originator. But, in this instance, it would have been unseemly for the ‘A’ and the ‘H’ authors to suggest such an order; the proposal would have had to come from the ‘W’ author. When this did not happen, the chivalrous suggestion was made that as Wu was the only woman she might sign first. (One wonders whether 40 years on such a suggestion would be regarded as an early example of affirmative action or a sexist remark!) So the authors of the paper describing the NBS parity violation experiment were listed as Wu, Columbia University, and underneath, on a separate line, Ambler, Hayward, Hoppes and Hudson, NBS³.

This gave the impression that Wu was the principal author, and as a result the experiment was and is often referred to as the “Wu experiment”. This attribution was first made by Garwin *et al.*⁴, whose paper on parity violation in meson decay followed the paper by Wu *et al.*³. They thanked Wu for “reports of her results of the ^{60}Co experiments”, rather than the NBS results. When in 1978 Wu was chosen as the first physics recipient of the prestigious Wolf Prize, the the NBS experiment was cited as “her most famous work”.

The purpose of this note is to state for the record that the NBS parity violation experiment was a collaborative team effort in which nuclear physicists and cryophysicists pooled their knowledge and expertise to carry out an experiment proposed by Lee and Yang, thus confirming their hypothesis that parity is not conserved in β -decay. In writing this article, we relied on published papers, unpublished or personal information and somewhat shaky reminiscences of one of us (N. K.). We are grateful to Ralph Hudson for putting us right on some facts. □

Nicholas Kurti and Christine Sutton are at the Physics Department of the University of Oxford, Oxford OX1 3RH, UK.

1. Lee, T. D. & Yang, C. N. *Phys. Rev.* **104**, 254 (1956).
2. Blackett, P. M. S. *Proc. R. Soc.* **251**, 301 (1959).
3. Wu, C. S., Ambler, E., Hayward, R. W., Hoppes, D. D. & Hudson, R. P. *Phys. Rev.* **105**, 1413 (1957).
4. Garwin, R. L., Lederman, L. M. & Weinrich, M. *Phys. Rev.* **105**, 1415 (1957).

An oblique slant on deep-sea biodiversity

Michael A. Rex

Organisms at the bottom of the deep seas might seem to be immune from large-scale shifts in climate. Not so, according to a study which reveals links between benthic biodiversity, glaciation and the Earth's obliquity.

Early theories on deep-sea biodiversity assumed that the sea floor is spatially and temporally uniform, and is buffered from climatic change at the surface by the overlying water column. That changed with Hessler and Sanders¹ discovery of surprisingly high species diversity in the deep sea, just 30 years ago, which inspired a period of intense ecological research in this largest and least known of Earth's ecosystems².

In consequence, the perspective of ecological stability and independence has gradually given way to recognition of a much more dynamic and heterogeneous environment, one which functions as an integral part of the global biosphere. For example, we now know that the structure of biological communities shows pronounced variation on local, regional and global scales^{3,4}, and that surface-benthic coupling is mediated through the rapidly sinking remains of phytoplankton (phytodetritus)^{5,6}, and through organismic life cycles that include migration between benthic and pelagic habitats⁷. But this new picture of a complex and variable deep sea has been restricted to brief ecological timescales — we have had only three decades of quantitative sampling, and *in situ* experiments have lasted only a year or two.

On page 624 of this issue⁸, Cronin and Raymo add an entirely new dimension to our understanding by showing that deep-sea biodiversity can also vary on geological timescales of 10^3 – 10^4 years. These long-term fluctuations in diversity correspond in a remarkably consistent way to climatic shifts at the surface associated with glaciation and ultimately linked to changes in solar insolation that attend the planet's 41,000-year obliquity cycle. During this cycle, the tilt of the Earth's axis of rotation varies with respect to the plane of its orbit, altering seasonality.

Cronin and Raymo's study focuses on benthic ostracod species, minute crustaceans from a core of the ocean floor taken in the North Atlantic by the Deep Sea Drilling Project. Ostracods were sampled

along the core at intervals representing 3,500–4,000 years; altogether, they span a 450,000-year segment of the Late Pliocene, 2.85 to 2.40 million years ago, during which well-documented glacial–interglacial cycles occurred. Ostracods are preserved in deep-sea sediments because of their heavily calcified bivalved shells, and they are the only deep-sea multicelled organisms to have left an abundant fossil record. Variation in their species make-up along the core provides a natural experiment of how deep-sea diversity changes with time in response to glaciation.

Cronin and Raymo were able to analyse this relationship in a precise way. Their approach was to compare ostracod species diversity to oxygen-isotope ratios in the remains of foraminiferans (shelled protozoans, which were collected from the same position in the core, and so have the same geological age), and to Mg:Ca ratios in the ostracod valves. These ratios reflect the history of surface-ice volume and bottom-

water temperatures, respectively. The relationships between ostracod species diversity and these climatic indicators were followed through 11 cycles of glaciation. The results show that diversity is depressed during glacial advances and recovers during interglacial phases throughout the time sampled. Not only do climatic and diversity cycles match with a high degree of statistical significance, but the amplitudes of the changes also correspond. Diversity declines most during severe glacial episodes.

Cronin and his colleagues show elsewhere^{9,10} that these glacial oscillations in diversity are not accompanied by extinction and origination of species. Rather, they represent a periodic attenuation of an assemblage of organisms that persists through the time sampled. Data from other deep-sea cores reveal that some species lost to the assemblage during glacial advances have shallower (but still deep-sea) refugia from which they can later recolonize affected areas. So, at least on geological timescales of 10^5 years in this phase of the Pliocene, local ostracod diversity fluctuates, and evolutionary diversification is not apparent in this region of the North Atlantic. Along with related evidence on foraminiferans from earlier in the Cenozoic^{11,12}, some 30–60 million years ago, Cronin and Raymo's findings require a re-evaluation of the basic notions that physical and biotic stability in the deep sea on geological timescales have promoted diversity, and that deep-sea communities are immune from major climatic events.

What is the causal connection between glaciation and deep-sea ostracod diversity? Cronin and Raymo doubt that direct effects of variation in bottom-water temperatures

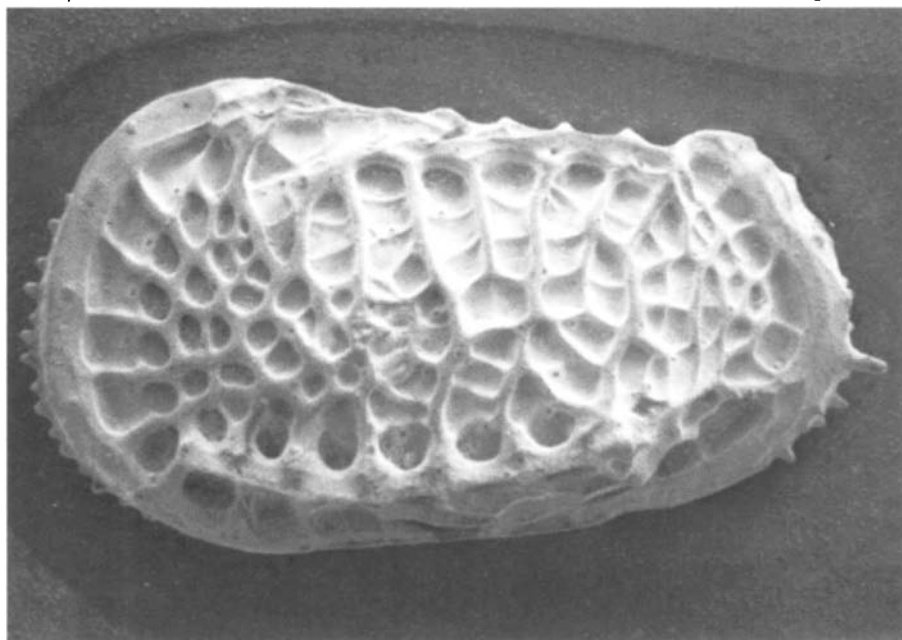


Figure 1 Scanning electron micrograph of *Poseidon amicus*, a member of deep-sea ostracod assemblages that are characteristic of interglacial periods. In reality, the length of this left valve is about 1 mm. (Micrograph courtesy of Deena Grinbaum.)

and dissolved nutrients are involved. Instead, they propose that the answer lies in climatically driven shifts in surface production and consequently nutrient input to the benthos from sinking phytodetritus. As a generality, this seems quite reasonable, because community structure in living deep-sea assemblages has been associated with nutrient input, and there is evidence of variation in overhead production from pelagic elements found in deep-sea cores.

Beyond this, it is hard to be more specific. The relationship between nutrient input (or productivity in surface environments) and species diversity has become a hot issue in macroecology. But there is still little agreement on the exact biological interactions involved, the scales on which they might operate, whether the mechanisms are equilibrium or nonequilibrium, and, indeed, even the correct statistical shape of the relationship¹³⁻¹⁶. Cronin and Raymo understandably steer clear of this empirical and theoretical quagmire, which can be approached critically only through controlled experimental tests on living deep-sea communities, and then with great difficulty. From their results, however, it seems increasingly likely that one or more features of nutrient input (rate, spatio-temporal variation, quality) are involved in shaping patterns of deep-sea biodiversity, past and present.

The sheer scale of events invoked in this study is certainly unprecedented in deep-sea ecology. Cronin and Raymo draw plausible links between diversity in the deep-sea benthos through geological time, surface production, patterns of glaciation and planetary behaviour. Their analysis shows the tremendous promise of a multidisciplinary approach, including palaeontology, palaeoclimatology and contemporary ecology, to develop a historical, global-scale understanding of deep-sea biodiversity. □

Michael A. Rex is in the Department of Biology, University of Massachusetts, 100 Morrissey Boulevard, Boston, Massachusetts 02125, USA.

- Hessler, R. R. & Sanders, H. L. *Deep-Sea Res.* **14**, 65-78 (1967).
- Gage, J. D. & Tyler, P. A. *Deep-Sea Biology: A Natural History of Organisms at the Deep-Sea Floor* (Cambridge Univ. Press, 1991).
- Etter, R. J. & Grassle, J. F. *Nature* **360**, 576-578 (1992).
- Rex, M. A. *et al.* *Nature* **365**, 636-639 (1993).
- Billett, D. S. M., Lampitt, R. S., Rice, A. L. & Mantoura, R. F. C. *Nature* **302**, 520-522 (1983).
- Graf, G. *Nature* **341**, 437-439 (1989).
- Young, C. M. & Eckelbarger, K. J. (eds) *Reproduction, Larval Biology, and Recruitment of the Deep-Sea Benthos* (Columbia Univ. Press, New York, 1994).
- Cronin, T. M. & Raymo, M. E. *Nature* **385**, 624-627 (1997).
- Cronin, T. M., Raymo, M. E. & Kyle, K. P. *Geology* **24**, 695-698 (1996).
- Cronin, T. M. *et al.* *Paleoceanography* **10**, 259-281 (1995).
- Kennett, J. P. & Stott, L. D. *Nature* **353**, 225-229 (1991).
- Thomas, E. & Gooday, A. J. *Geology* **24**, 355-358 (1996).
- Ricklefs, R. E. & Schluter, D. (eds) *Species Diversity in Ecological Communities: Historical and Geographical Perspectives* (Univ. Chicago Press, 1993).
- Huston, M. A. *Biological Diversity: The Coexistence of Species on Changing Landscapes* (Cambridge Univ. Press, 1994).
- Brown, J. H. *Macroecology* (Univ. Chicago Press, 1995).
- Rosenzweig, M. L. *Species Diversity in Space and Time* (Cambridge Univ. Press, 1995).

Schizophrenia

An emerging pathophysiology

Eric J. Nestler

The aetiology and pathophysiology of schizophrenia have been the focus of intensive research for decades, during which time the so-called 'dopamine hypothesis' has predominated. Because all clinically effective antipsychotic drugs are antagonists of D2-like dopamine receptors (which include the D2-, D3- and D4-receptor subtypes), excessive dopamine-mediated neurotransmission or high levels of D2-like receptors in the brain were thought to underlie schizophrenia.

In recent years, the dopamine hypothesis has given way to more complicated — and realistic — views of the disorder. The new picture invokes pathophysiological mechanisms in specific brain pathways that involve more than a simple increase or decrease in the activity of a single neurotransmitter or receptor type. It is in this context of an emerging, and more sophisticated, view of schizophrenia, that Okubo *et al.*¹ report their findings on page 634 of this issue. They have used positron emission tomography to show that there are reduced levels of D1-like dopamine receptors (which include the D1- and D5-receptor subtypes) in the prefrontal cortex of patients with schizophrenia. This reduction is also seen in patients who have never been exposed to antipsychotic drugs — an important control, as chronic treatment with antipsychotic drugs can also decrease levels of the D1 receptor in the prefrontal cortex of non-human primates².

This is an exciting time for research into schizophrenia. The development of brain imaging has allowed the brains of schizophrenics to be studied directly, and it is becoming increasingly feasible to understand these images within a cellular and molecular context. Considerable attention has been paid to the prefrontal cortex which, along with related limbic cortical structures (for example, the cingulate cortex), has been implicated as an important neural substrate of schizophrenia. These regions are thought to contribute particularly to the cognitive impairment and 'negative symptoms' (decreased motivation and expressed emotion) of schizophrenia; but they may also contribute to the 'positive symptoms', such as hallucinations and delusions.

Much is known about the expression and cellular/subcellular localization of dopamine receptors within the primate prefrontal cortex. The expression of D1-like receptors is around ten times higher than that of D2-like receptors. The D1 receptors are expressed predominantly by pyramidal neurons, on their dendritic spines³ (Fig. 1). This is an ideal location for modulation of the glutamate-mediated inputs to these neurons — inputs that mainly come from other pyramidal neurons and thalamic neurons. The D5 receptors are also expressed (at lower levels) by pyramidal neurons, but they are localized to the shafts of dendrites⁴. The distribution of D2-like receptors is not as

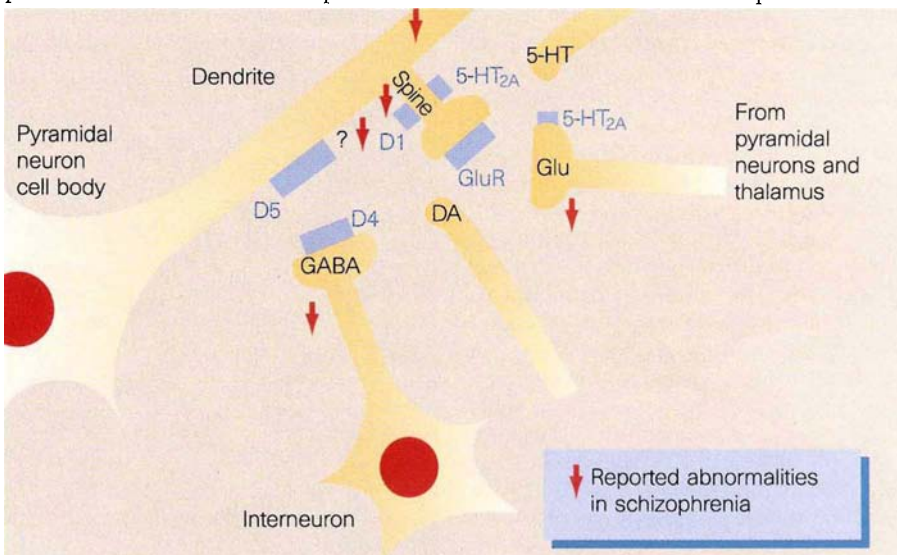


Figure 1 Localization of dopamine receptors within the synaptic organization of the prefrontal cortex (highly simplified). Okubo *et al.*¹ have found that one of the neural abnormalities associated with schizophrenia is a reduction in the levels of D1-like (D1 and D5) receptors. D1 receptors are restricted to dendritic spines, close to glutamate receptors; D5 receptors are localized to dendritic shafts; D4 receptors are enriched in GABA-containing interneurons; D2 and D3 receptors are present at lower levels in the prefrontal cortex, and they may be expressed in several neuronal elements. Glu, glutamate; GluR, glutamate receptors; GABA, γ -aminobutyric acid; DA, dopamine; 5-HT, serotonin.

completely understood, but they seem to be expressed in pyramidal cells as well as in γ -aminobutyric acid (GABA)-containing interneurons (local-circuit neurons) and glutamate-containing nerve terminals that innervate the region⁵.

Studies in non-human primates have begun to provide information about the functional consequences of dopamine-mediated neurotransmission in the prefrontal cortex. One view is that certain pyramidal neurons within the cortex form the cellular basis for working memory, and that dopamine can modulate this process largely through D1-like receptors⁶. There seems to be an optimal level of D1-like receptor activation, above or below which working memory is impaired, and similar mechanisms may operate in the cingulate cortex⁷.

The reduced levels of D1-like receptors seen by Okubo *et al.*¹ in the prefrontal cortex of schizophrenic patients could, therefore, contribute to deficiencies in certain aspects of cognition. This is consistent with their observation that the lower the levels of D1-like receptors, the more severe the negative symptoms. On the other hand, antipsychotic drugs themselves decrease D1-receptor levels², so one must ask whether this drug-induced decrease mediates therapeutic or side effects of the drugs. And is the observed reduction in D1-receptor levels in schizophrenics part of the pathology of the disorder or a compensatory adaptation to that pathology?

Autopsy studies have also focused on cellular abnormalities in the prefrontal and cingulate cortices of schizophrenics. Variable results have been reported: a decrease in the number of GABA-containing interneurons⁸; a reduction in the amount of GABA made by these neurons, but no loss in their number⁹; a reduction in the dendritic processes of pyramidal neurons¹⁰; reductions in the levels of various presynaptic-nerve-terminal markers¹¹; and increased levels of nerve-cell bodies in white matter, which suggests a defect in neuronal migration¹². These discrepancies emphasize the need for further research. For example, it will be essential to determine whether schizophrenia involves the loss of cortical neurons and their inputs or, rather, a sustained impairment in the activity of existing neurons.

These autopsy studies are nonetheless encouraging because, for the first time, we have a lead as to the cellular pathophysiology of schizophrenia. The findings can also be reconciled with brain-imaging studies of the disorder. In general, autopsy studies point to a reduction in neuronal mass or activity, which could relate to the smaller volume of grey matter and reduced blood flow that are seen in the frontal cortex of schizophrenics^{13,14}. The low levels of D1-like receptors reported by Okubo *et al.*¹ could reflect the reduction in dendritic processes

of pyramidal neurons¹⁰ (where the D1-like receptors are enriched³).

Nevertheless, it is critical to emphasize that many neurotransmitter and receptor types are probably involved in schizophrenia. For example, the serotonin 5-HT_{2A} receptors¹⁵, which are localized to the distal dendrites of pyramidal neurons and some glutamate-containing nerve terminals (Fig. 1), modulate glutamate-mediated transmission. Hallucinogens — which mimic some of the positive symptoms of schizophrenia — are partial agonists at the 5-HT_{2A} receptors, whereas many antipsychotic drugs, particularly the 'atypical' drugs (defined by their improved efficacy and fewer side effects), are antagonists at these receptors.

One of the main questions in the field concerns the nature of the original insult that leads to cellular abnormalities in the prefrontal and cingulate cortices. Genetic and epidemiological studies indicate that there is a strong genetic component to the disorder, and that unidentified environmental factors are also required. Autopsy studies agree on the fact that there is no gliosis, inflammation or active neural degeneration, indicating that schizophrenia could involve an earlier insult during which time the damage is done. Alternatively, schizophrenia could involve a continuing pathophysiological process that impedes the functional activity of certain cortical neurons, but without obvious neurodegeneration. Reduced activity would be consistent with the clinical observation that some schizophrenics with severe chronic symptoms show dramatic improvement upon treatment with atypical drugs such as

clozapine. If this model holds, it clearly offers much hope to schizophrenics as it should allow the development of more effective treatments.

Finally, it is important to emphasize that schizophrenia is probably a heterogeneous disorder, with several aetiologies resulting in a syndrome that has some shared clinical features. Brain-imaging and autopsy studies are based on the idea that, as well as some shared clinical features, there may be shared pathophysiological mechanisms. However, in medicine, the precedent is that similar-looking physiological and behavioural defects arise from distinct molecular and cellular abnormalities. In any event, sophisticated neurobiological hypotheses of schizophrenia are now emerging, which define the critical directions for future research. □

Eric J. Nestler is in the Laboratory of Molecular Psychiatry, Yale University School of Medicine, New Haven, Connecticut 06508, USA.

1. Okubo, Y., Suhara, T. & Suzuki, K. *Nature* **385**, 634–636 (1997).
2. Lidow, M. S. & Goldman-Rakic, P. S. *Proc. Natl Acad. Sci. USA* **91**, 4353–4356 (1994).
3. Smiley, J. F. *et al. Proc. Natl Acad. Sci. USA* **91**, 5720–5724 (1994).
4. Bergson, C. *et al. J. Neurosci.* **15**, 7821–7836 (1995).
5. Mrzljak, L. *et al. Nature* **381**, 245–248 (1996).
6. Williams, G. V. & Goldman-Rakic, P. S. *Nature* **376**, 572–575 (1995).
7. Dolan, R. J. *et al. Nature* **378**, 180–182 (1995).
8. Benes, F. M. *Arch. Gen. Psychiatry* **52**, 1015–1018 (1995).
9. Akbarian, S. *et al. Arch. Gen. Psychiatry* **52**, 258–266 (1995).
10. Selemon, L. D., Rajkowska, G. & Goldman-Rakic, P. S. *Arch. Gen. Psychiatry* **52**, 805–818 (1995).
11. Eastwood, S. L. & Harrison, P. J. *Neuroscience* **69**, 339–343 (1995).
12. Akbarian, S. *et al. Arch. Gen. Psychiatry* **53**, 425–436 (1996).
13. Kotlir, K. J. & Weinberger, D. R. *Annu. Rev. Med.* **46**, 113–122 (1995).
14. Andreasen, N. C. *Neuron* **16**, 697–700 (1996).
15. Aghajanian, G. K. & Marek, G. J. *Neuropharmacology* (in the press).

Protein evolution

How far can sequences diverge?

Cyrus Chothia and Mark Gerstein

Protein interiors consist of closely packed amino-acid residues. Close packing places strong constraints on side-chain positions, and means that the large sequence changes that occur in some protein families over long periods of evolution lead to considerable structural variations — α -helices and β -sheets shift relative to each other, and loop regions change conformation¹. These types of modifications are acceptable if they maintain the stability of the protein and do not adversely affect its function. Structural changes need not affect function if they are distant from the active site; and if they are close to the active site then they can be coupled in such a way that function is conserved². But how do the conformational requirements of function limit the extent to which proteins can alter their sequences during evolution? Two experiments in protein engineering, reported by Gassner *et al.*³ and by Axe *et al.*⁴ in *Proceed-*

ings of the National Academy of Sciences, provide results that are important for answering this question.

Gassner *et al.*³ studied T4 lysozyme, which has an active site in a groove between two domains (Fig. 1, overleaf). They replaced ten residues that are buried in the core of the second all-helical domain (six leucines, two isoleucines, one valine and one phenylalanine) with methionine residues. These mutations make little or no change to the volumes of the side chains that they replace, but they do change their shape. They found that the individual mutations had small effects on function — the mutant proteins had between 70% and 105% of the native activity, and a loss in stability of 0.4–0.9 kcal energy per mol of protein. But a protein with seven of the mutations showed 43% of the native activity and a 5.0-kcal loss in stability.

By comparing the atomic structure of the

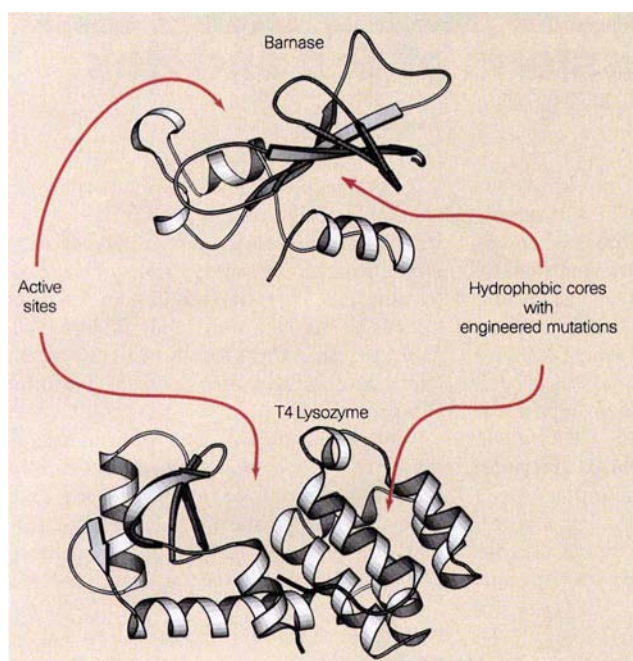


Figure 1 The molecular structure of barnase⁹ (top) and T4 lysozyme¹⁰ (bottom). Coiled ribbons represent α -helices and extended ribbons represent strands of β -sheet. Gassner *et al.*³ engineered mutations in T4 lysozyme at the sites of ten residues that pack between the α -helices that form the second domain, and Axe *et al.*⁴ engineered mutations in barnase at the sites of 13 residues that pack between the amino-terminal α -helix and the β -sheet. The picture was drawn using MOLSCRIPT¹¹.

T4 lysozyme containing seven mutations with that of the native protein, Gassner *et al.* were able to show how the protein had responded to the mutations. Six of the mutations involved the replacement of leucine by methionine, the side chains of which have the same volume but different shapes: whereas the leucine side chain is branched ($-\text{CH}_2-\text{CH}(\text{CH}_3)_2$), that of methionine is linear ($-\text{CH}_2-\text{CH}_2-\text{S}-\text{CH}_3$). The authors showed that the first three side-chain groups in the methionine residues ($-\text{CH}_2-\text{CH}_2-\text{S}-$) occupied the same positions as three of the side-chain groups ($-\text{CH}_2-\text{CH}-\text{CH}_3$) in the leucines that they replaced. However, the different stereochemistries of methionine and leucine meant that the terminal methyl groups in the six methionine residues occupied quite different positions (2.6–3.9 Å) from the branched methyl groups of the six leucines.

Gassner *et al.*³ next superposed the helical domains of the mutant and native proteins, and they found that the overall root-mean-square difference in the position of the main-chain atoms between the two structures is 0.2 Å. They suggest that this close similarity in the structures of the two proteins is not compatible with the precise spatial complementarity that is implied by the close-packed description of the protein interior.

Stimulated by these findings, we analysed the mutant and native structures. Although the overall difference in the position of the main-chain atoms is small, we find that there are significant local changes. In particular, the main-chain atoms of two regions that contain a mutation (residues 82–84 and 153–155) shift by 0.4–0.7 Å. Also, in a region that is adjacent to one mutation and in contact with two others, one helix and an adjacent turn (residues 107–116) are slightly distorted, and they shift by 0.4–1.0 Å. Buried side chains that are adjacent to the mutations

rotate by 10–23°, and an adjacent surface side chain rotates by 58°.

We also determined the volumes that are occupied by the buried atoms of the mutant methionine residues. Two-thirds of the atoms in the six methionines are totally buried, and they occupy a total of 612 Å³. The volume that would be expected for these atoms — if they are close-packed in the usual manner⁵ — is 614 Å³. (This volume implies a greater packing density than that in amino-acid crystals⁵.) So the effect of the structural changes is to go from a set of close-packed leucines in the native structure to a set of close-packed methionines in the mutant. In other words, by using a variety of subtle changes in its conformation, T4 lysozyme can accommodate simultaneous changes in the shape of buried residues, and maintain close packing.

Axe *et al.*⁴ carried out a much more extensive series of mutations to the main hydrophobic core of barnase, a bacterial ribonuclease. The major secondary structure in this protein is a six-stranded β -sheet, one face of which, together with certain loop regions, forms the active site. A helix packs against the other face of the β -sheet, and the region between these two secondary structures forms the main hydrophobic core of the protein (Fig. 1).

The authors made random hydrophobic substitutions at the 13 residues that form the hydrophobic core. In the native structure, these residues are: two alanines, one valine, four isoleucines, two leucines, one phenylalanine, two tyrosines and one tryptophan. They were randomly replaced by valine, leucine, isoleucine, methionine or phenylalanine, and Axe *et al.* found that a surprisingly high proportion of the mutants — close to a quarter — had enzymatic activity. Even in some cases where all 13 sites had non-native residues, mutants were highly active.

The experiments by Axe *et al.*⁴ and Gassner *et al.*³, together with previous studies by Lim *et al.*^{6,7}, imply that protein structures can often adapt to changes in the shape of buried residues, and that many different combinations of hydrophobic residues in the core are able to maintain structural integrity.

Gassner *et al.*³ have shown that the presence of seven mutations within the active site of T4 lysozyme reduces the activity by half, even though the structural changes are small. When single mutations are introduced, we would expect the structural changes to be even smaller, yet function is again affected, with activities from 70–105% of that of the native protein. The mutations introduced by Axe *et al.*⁴ affect the activity of barnase to a lesser degree and the mutant proteins are highly active, even though the mutations must produce changes that are much larger than those in T4 lysozyme. This must be because the active site is on the opposite side of the β -sheet to the mutations, and it will only be affected if conformational changes are transmitted across the β -sheet.

Conformational changes allow individual structures to accept a range of mutations, especially those that retain the chemical character of the residue and do not involve large changes in volume in one step. Limitations on the extent of this process in actual proteins arise from their functional requirements. For proteins such as histones and actin, whose functions involve extensive interactions with several different proteins or DNA, there are severe limitations — the rate of change in their sequences is about 10% per billion years⁸. Conversely, for the different domains in the immunoglobulin superfamily, where duplicated genes have different functions, less than a third of the domain structure is conserved, and the only sequence conservation is the hydrophobic character and approximate size of some 12 residues. Essentially, the extent of sequence divergence in proteins is inversely related to the extent of their functional requirements. □

Cyrus Chothia is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. Mark Gerstein is in the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114, USA.

1. Chothia, C. & Lesk, A. M. *EMBO J.* **5**, 823–826 (1986).
2. Lesk, A. M. & Chothia, C. *J. Mol. Biol.* **136**, 223–268 (1980).
3. Gassner, N. C., Baase, W. A. & Matthews, B. W. *Proc. Natl. Acad. Sci. USA* **93**, 12155–12158 (1996).
4. Axe, D. D., Foster, N. W. & Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **93**, 5590–5594 (1996).
5. Harpaz, Y., Gerstein, M. & Chothia, C. *Structure* **2**, 641–649 (1994).
6. Lim, W. A. & Sauer, R. T. *Nature* **339**, 31–36 (1989).
7. Lim, W. A., Hodel, A., Sauer, R. T. & Richards, F. M. *Proc. Natl. Acad. Sci. USA* **91**, 423–427 (1994).
8. Doolittle, R. F. *Protein Sci.* **1**, 191–200 (1992).
9. Mauguén, Y. *et al. Nature* **297**, 162–164 (1982).
10. Weaver, L. H. & Matthews, B. M. *J. Mol. Biol.* **193**, 189–199 (1987).
11. Kraulis, P. J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).

NATURE

100 YEARS AGO

In consequence of my measurements of Kerr's magneto-optical phenomena, the thought occurred to me whether the period of the light emitted by a flame might be altered when the flame was acted upon by magnetic force. It has turned out that such an action really occurs....The possibility of an alteration of period was first suggested to me by the consideration of the accelerating and retarding forces between the atoms and Maxwell's molecular vortices; later came an example suggested by Lord Kelvin, of the combination of a quickly rotating system and a double pendulum. However, a true explanation appears to me to be afforded by the theory of electric phenomena propounded by Prof. Lorentz. In this theory, it is considered that, in all bodies, there occur small molecular elements charged with electricity, and that all electrical processes are to be referred to the equilibrium or motion of these "ions". It seems to me that in the magnetic field the forces directly acting on the ions suffice for the explanation of the phenomena. Professor Lorentz ...was good enough to show me how the motion of the ions might be calculated, and further suggested that if my application of the theory be correct there would follow these further consequences: that the light from edges of the widened lines should be circularly polarised when the direction of vision lay along the lines of force; further, that the magnitude of the effect would lead to the determination of the ratio of the electric charge the ion bears to its mass.... An altogether rough measurement gives 10^7 as the order of magnitude of the ratio e/m when e is expressed in electromagnetic units. "The effect of magnetisation on the nature of light emitted by a substance" — P. Zeeman. From *Nature* 11 February 1897.

50 YEARS AGO

The extent of the malnutrition can be estimated from the finding that more than half the world's population had a pre-war calorie intake of only 2,250 a head a day; less than one third had available 2,750 or more calories.... In some of the countries in the low-calorie group, such as Java, the malnutrition was worse than the total figures suggest, since the proportion of cereals consumed was too high and the protein and fat consumption quite inadequate to maintain health. From *Nature* 15 February 1947.

Protein-tyrosine kinases

New impressions of Src and Hck

Tony Pawson

The Src cytoplasmic tyrosine kinase has captured the imagination and experimental attention of those scientists interested in malignant transformation and signal transduction for most of this century. Yet until now, we have only had an impressionist's view of Src — the general picture has been perceived but the details have been fuzzy. This view changes with the reports on pages 595 and 602 of this issue, by Xu *et al.*¹ and Sicheri *et al.*², who have solved the crystal structures of human c-Src and its close relative Hck (haematopoietic cell kinase), respectively. In each case, the kinase has been inactivated by the phosphorylation of

an inhibitory carboxy-terminal tyrosine residue (tyrosine residue 527 or its equivalent). In a related paper on page 650, Moarefi *et al.*³ show that the kinase activity of Hck can be stimulated by the binding of specific ligands to the Hck non-catalytic domains. Taken together, these results bring the regulatory mechanisms of c-Src and Hck into sharp focus.

The oncogenic properties of Rous sarcoma virus⁴ — an avian retrovirus that causes the rapid eruption of tumours in infected animals — led to the identification of the viral oncogene *v-src*, and to the recognition that this is a subtly altered form of the normal

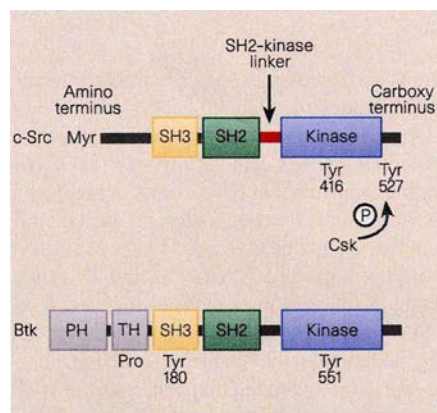


Figure 1 The domain organization of c-Src and Btk. The c-Src protein is anchored to the cell membrane through a myristylation (Myr) site, whereas Btk is potentially attached through a pleckstrin homology (PH) domain. Both proteins contain Src-homology (SH) domains 2 and 3, along with a catalytic kinase domain. The Tec-homology (TH) domain of Btk contains a proline-rich motif. The activity of c-Src is controlled by phosphorylation at Tyr 527, and the structural analyses of Xu *et al.*¹ and Sicheri *et al.*², along with the biochemical data of Moarefi *et al.*³, point to the mechanism by which this may occur.

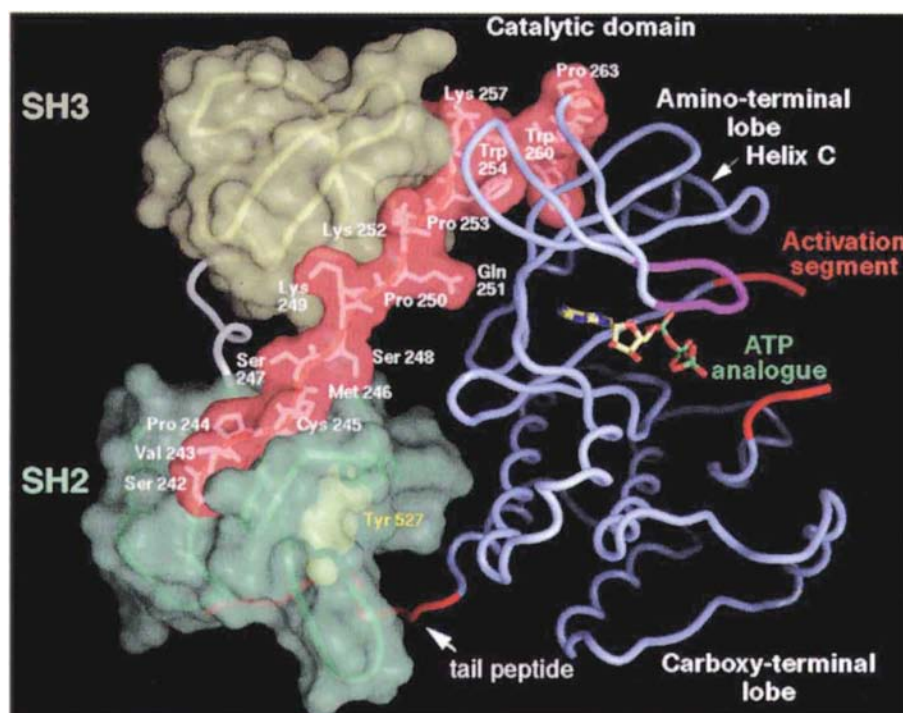


Figure 2 Structure of Hck. The molecular surfaces of the SH3 domain (yellow) and SH2 domain (green), and the SH2-kinase linker (red) are superimposed on a backbone representation of inactive Hck. The backbone of the kinase domain is shown in blue. (The figure was provided by F. Sicheri, I. Moarefi and J. Kuriyan.)

cellular gene, *c-src*⁵. Although the over-expression of unadulterated c-Src has little effect on cellular behaviour, the substitution of a single amino acid can convert the protein to a form that induces profound changes in cell growth control, gene expression, metabolism and cytoskeletal architecture.

Why do oncogenic variants of c-Src have such dire effects on cellular behaviour? A key finding, in this regard, was that Src is a protein kinase^{6,7}, which specifically phosphorylates tyrosine residues⁸. The enzymatic activity of c-Src is normally very low, but is constitutively elevated in transforming variants, meaning that they can aberrantly activate signalling pathways that are controlled by tyrosine phosphorylation. This raises two important questions: how is the c-Src kinase activity normally repressed, and how does c-Src interact with its cellular targets once this inhibition is released? The answers are closely connected, and revolve around the idea that signalling by tyrosine kinases is regulated by modular protein-protein interactions⁹.

The c-Src protein is the prototype for a family of nine tyrosine kinases — including Hck — all of which are structurally related and are often involved in signalling from antigen and cytokine receptors. The amino terminus of c-Src is modified by myristylation (covalent attachment of fatty acid), which is required both for membrane association and biological activity. All members of the Src family possess a variable region followed by SH3 and SH2 domains, the catalytic (kinase) domain, and a short carboxy-terminal tail (Fig. 1).

SH2 and SH3 domains mediate the formation of specific protein complexes⁹, and are found in a range of cytoplasmic signalling proteins. SH2 domains have two binding pockets, one for phosphotyrosine, and a second, more variable site that determines specificity by binding the residues on the carboxy-terminal side of the phosphotyrosine. SH3 domains generally recognize motifs containing the sequence PXXP, which adopt a polyproline 'type-II' (PP-II) helix. Usually, SH3 domains comprise five β -strands that form a hydrophobic binding surface for the proline-rich sequence, which is flanked by two connecting loops (the 'RT' and 'n-Src' loops). These loops can also make specific contacts with ligands, thereby increasing the affinity of the interaction¹⁰.

The activity of c-Src is mainly controlled by its carboxy-terminal tail, which possesses a tyrosine residue — Tyr 527 of chicken c-Src — that can be phosphorylated by the cytoplasmic tyrosine kinase, Csk. Counter-intuitively perhaps, phosphorylation of Tyr 527 represses the kinase activity of c-Src. Indeed, the main alteration in v-Src is a deletion that removes Tyr 527, and c-Src can be converted to an efficient oncogene by substituting Tyr 527 with phenylalanine.

Active galaxies

A safer hiding hole

Old quasars probably just fade away. According to the conventional theory, all that they, and other active galaxies, leave behind is a huge black hole sitting in an ordinary galaxy — one that lacks the brilliant emissions of quasars because no gas is falling into the black hole to act as gravitational fuel. There is dynamical evidence for these black holes in several galaxies, but as John Kormendy warns, "This situation is dangerous: it is easy to believe that we have proved what we expected to find" (*Annu. Rev. Astron. Astrophys.* 33, 581–624; 1995). So it is comforting that, in this issue, van der Marel *et al.* (*Nature* 385, 610–612; 1997) have improved one strong case.

Their analysis is based on new spectral observations of stars near the centre of M32, the small elliptical galaxy shown here. The high velocities of these stars demand that two to five million solar masses of dark material exist within a sphere one light year across. It is probably a single massive black hole. A cluster of smaller objects is unlikely: being a billion times denser than that of our stellar neighbourhood, either the cluster would rapidly collapse into a single central object by collisions and gravitational radiation, or its constituents would soon disintegrate into a bright gas cloud.

The same arguments apply to other

galaxies. NGC3115, for example, must hold a dark object of about two billion solar masses to keep its central star cluster from flying apart (J. Kormendy *et al. Astrophys. J. Lett.* 459, L57–L60; 1996). It seems likely that most galaxies have been quasar-like at some time.

But some galaxies, such as M33, do not contain supermassive black holes. Are they only an incidental part of galaxy evolution? Or are they essential, but liable to get lost in collisions with other black-hole-containing galaxies, and then wander dimly through intergalactic space? With a large enough survey of galaxies we should be able to answer these questions.

Stephen Battersby

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

One hypothesis to explain these findings is that the phosphorylated carboxy-terminal tail might forge an intramolecular interaction with the SH2 domain, and in doing so interfere with the intervening kinase domain¹¹ (Fig. 1). Consistent with this view, mutations that affect the ability of the c-Src SH2 domain to bind phosphotyrosine result in kinase activation. However, the ligand-binding surface of the c-Src SH3 domain is also essential for kinase inhibition, indicating that the SH3 domain may contact the kinase domain¹².

Protein-kinase catalytic domains are typically divided into a small ATP-binding lobe and a large peptide-binding lobe, with the active site located in the cleft between the two lobes. Xu *et al.*¹ and Sicheri *et al.*² have found that the catalytic domains of inactive c-Src and Hck have a similar structure to the catalytic domain of the Src-family member Lck, which has been crystallized in the active form¹³. In this active state, Tyr 527 is unmodified; however, Tyr 416 in the activation loop of the kinase domain is phosphorylated.

Xu *et al.* show that inactive c-Src (in which Tyr 527 is phosphorylated) has more closely apposed ATP- and peptide-binding lobes than are found in active kinase structures — as if they have been forced together following the phosphorylation. As anticipated, the

phosphorylated carboxy-terminal tail interacts with the SH2 domain (Fig. 2) but, surprisingly, this interaction occurs on the opposite side of the molecule from the kinase active-site, and it does not seem to disturb the structure of the kinase domain.

Without doubt, the most unexpected observation is that the peptide sequence that links the SH2 and kinase domains in both c-Src and Hck adopts a PP-II helix, even though c-Src does not contain a PXXP motif. The core ligand-binding surface of the SH3 domain recognizes this linker region, while regions of the RT loop (among them the Arg 95 and Thr 96 residues after which the loop is named) contact the small lobe of the kinase catalytic domain. So a key function of the interaction between phosphotyrosine and the SH2 domain may be to position the SH3 domain in such a way that it can bind the SH2-kinase linker, and consequently contact the small lobe of the kinase domain.

But how do these gymnastic contortions lead to inhibition of the kinase activity, given that they occur on the face of the kinase that is removed from the active site? The answer seems to lie in the position of an α -helix (helix C), in the small lobe of the kinase catalytic domain. In the inactive forms of c-Src or Hck, the SH3 domain and SH2-kinase linker push on one end of helix C, while the

activation loop from the large lobe butts into the other end. The resulting position of helix C forces a residue that is probably critical for kinase activity — Glu 310 — away from the active site. Activation of the kinase (for example, by dephosphorylation of the tail), would presumably lead to the sequential dissociation of the SH2 and SH3 domains from their intramolecular ligands, removing one constraint on helix C. Furthermore, autophosphorylation of Tyr 416 in the activation loop would result in reorganization of this sequence, thereby lifting the second constraint.

Why adopt this remarkably complex way of regulating Src-family kinases? Xu *et al.*¹ and Sicheri *et al.*² have shown that when Tyr 527 is phosphorylated, not only is the kinase domain repressed, but the SH2 and SH3 domains are tied up in intramolecular interactions, such that they cannot associate with exogenous ligands and disturb SH2/SH3-regulated signalling events. Does this mean that c-Src is activated simply by dephosphorylation of Tyr 527? Although this is likely to be the key regulatory switch, things may be more complicated. For example, the intramolecular interactions in which the SH2 and SH3 domains are engaged are probably very weak, and more potent exogenous ligands for either domain may compete for binding, thereby disrupting the inactive structure.

Moarefi *et al.*³ have addressed this issue by showing that a phosphopeptide that binds with high affinity to the Hck SH2 domain stimulates kinase activity (as has previously been shown for c-Src¹⁴), presumably by competing the SH2 domain away from phosphorylated Tyr 527. This may be of physiological relevance, as autophosphorylation of the β platelet-derived growth-factor receptor creates a binding site for the c-Src SH2 domain¹⁵, which could elevate the kinase activity by displacing the SH2 domain.

Moarefi *et al.* also show that the HIV Nef protein — a potent ligand for the Hck SH3 domain — stimulates the activity of Hck that is phosphorylated at Tyr 527, and even increases the activity of the dephosphorylated protein. So the SH3 domain could exert a residual repressive activity in the absence of the intramolecular phosphotyrosine-SH2 interaction. Several substrates for Src-family kinases have proline-rich motifs onto which the SH3 domain can dock, and this binding may tether the substrate to the kinase and concomitantly stimulate the kinase activity. Phosphorylation of a single tyrosine site in the substrate would then promote intermolecular SH2-binding, and hence a strong bidentate association between the enzyme and the substrate, which could stimulate processive phosphorylation of the substrate. This may be especially pertinent for substrates such as p130^{cas} and its relative Sin/Efs, which are phosphorylated at many sites by

Src-family kinases, and which once phosphorylated act as docking proteins for other signalling proteins with SH2 domains^{16,17}.

Modular interactions of the sort shown for Src kinases are also important for the Tec family of intracellular tyrosine kinases, among which the Btk kinase (Fig. 1) is well known for its involvement in human X-linked agammaglobulinaemia¹⁸. Btk has an amino-terminal pleckstrin homology (PH) domain that binds specifically to phosphatidylinositol-3,4,5-trisphosphate¹⁹, and may be involved in recruitment to the membrane. The PH domain is followed by a proline-rich sequence (the Tec homology, or TH, region), and by SH3, SH2 and kinase domains. Unlike the Src-family kinases, members of the Tec family are not regulated by carboxy-terminal phosphorylation: instead, the SH3 domain apparently undergoes an intramolecular interaction with the TH region, which could restrict access of the SH3 domain to potential substrates²⁰. Phosphorylation of a tyrosine residue within the SH3 domain²¹ may disrupt the intramolecular SH3-TH interaction, thereby liberating the SH3 domain to engage kinase substrates.

Tec-family kinases therefore depend on modular interactions for their regulation and target recognition — but in a rather different way from the Src-family kinases. Taken with the results of Xu *et al.*¹, Sicheri *et al.*² and Moarefi *et al.*³, these findings underscore the extraordinary versatility of protein modules such as SH2, SH3 and PH domains; however, there is clearly still much to be discovered about the structures, biochemistry and biological functions of Src and Tec kinases. □

Tony Pawson is at the Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, Canada M5G 1X5.

- Xu, W., Harrison, S. C. & Eck, M. J. *Nature* **385**, 595–602 (1997).
- Sicheri, F., Moarefi, I. & Kuriyan, J. *Nature* **385**, 602–609 (1997).
- Moarefi, I. *et al.* *Nature* **385**, 650–653 (1997).
- Rous, P. *J. Exp. Med.* **13**, 397–411 (1911).
- Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. *Nature* **260**, 170–173 (1976).
- Collett, M. S. & Erikson, R. L. *Proc. Natl Acad. Sci. USA* **75**, 2021–2024 (1978).
- Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. M. *Cell* **15**, 561–572 (1978).
- Hunter, T. & Sefton, B. M. *Proc. Natl Acad. Sci. USA* **77**, 1311–1315 (1980).
- Pawson, T. *Nature* **373**, 573–580 (1995).
- Lee, C.-H., Saksela, K., Mirza, U. A., Chait, B. T. & Kuriyan, J. *Cell* **85**, 931–942 (1996).
- Matsuda, M., Mayer, B. J., Fukui, Y. & Hanafusa, H. *Science* **248**, 1537–1539 (1990).
- Erpel, T., Superti-Furga, G. & Courtneidge, S. A. *EMBO J.* **14**, 963–975 (1995).
- Yamaguchi, H. & Hendrickson, W. A. *Nature* **384**, 484–489 (1996).
- Liu, X. *et al.* *Oncogene* **8**, 1119–1126 (1993).
- Mori, S. *et al.* *EMBO J.* **12**, 2257–2264 (1993).
- Sakai, R. *et al.* *EMBO J.* **13**, 3748–3756 (1994).
- Alexandropoulos, K. & Baltimore, D. *Genes Dev.* **10**, 1341–1355 (1996).
- Sideras, P. & Smith, C. I. E. *Adv. Immunol.* **59**, 135–223 (1995).
- Salim, K. *et al.* *EMBO J.* **15**, 6241–6250 (1996).
- Andreotti, A. H., Bunnell, S. C., Feng, G., Berg, L. J. & Schreiber, S. L. *Nature* **385**, 93–97 (1997).
- Park, H. *et al.* *Immunity* **4**, 515–525 (1996).

Daedalus

A problem dissolved

In popular medical demonology, cholesterol has a prominent place. It gets itself deposited in our arteries, often as local and dangerous constrictions.

Many epidemiologists blame components of our diet. Cholesterol itself is an obvious suspect; but since the body makes it as required and disposes of any surplus, it cannot be the whole story. Another suspect, alcohol, attracts both blame and praise. Countries such as France combine a high alcohol intake with relatively little heart disease. Can the antioxidants in red wine, or even alcohol itself, have some protective effect? Daedalus now points out the obvious protective mechanism. Cholesterol is slightly soluble in alcohol.

So, says Daedalus, a bloodstream loaded with alcohol should slowly dissolve cholesterol deposits, or should fail to lay them down in the first place. For best effect, the subject should drink steadily all the time, maintaining an even level of alcohol in his blood. This 'French' style of alcoholism contrasts strongly with the 'English' style of occasional wild alcoholic binges interspersed with periods of remorseful sobriety. While DREADCO's epidemiologists study the cardiological records of various styles of alcoholic, the firm's chemists are seeking a better inebriating solvent than alcohol.

One obvious alternative is ether. It dissolves cholesterol and its esters much better than alcohol, and is usefully soluble in blood. It is not very toxic (it is a traditional anaesthetic), and even causes its own brand of drunkenness, the nineteenth century 'ether frolic'. Many people might love to be kept slightly blotto on ether for several weeks, while the stuff strips the cholesterol from their arterial constrictions, and carries it to the liver for disposal — or simply redeposits it more evenly around the whole vascular system, without dangerous concentrations.

Another alternative is methyl *t*-butyl ether, MTBE. Though better known as an antiknock additive for petrol, it has in fact been used to dissolve cholesterol gall stones *in vivo*. It is not miscible with blood, and would have to be injected in emulsion form. MTBE might function as a sort of vascular flushing oil, stripping the deposits from the patient's arteries and holding them in suspension. Some sort of kidney-machine could filter off the suspended cholesterol and return the purified blood to the patient. Most people should prefer the occasional discomfort of flushing out to the grimly permanent rigours of a healthy lifestyle.

David Jones

Melvin Calvin (1911–97)

Chemist who elucidated the biochemical basis of photosynthesis

To achieve great things, being a great man is not enough: you also need to be in the right place at the right time. Melvin Calvin was lucky — although he was certainly an outstanding scientist, he had the good fortune to be in Berkeley at the end of the Second World War and to be a colleague of Ernest Lawrence.

As a boy helping in the family grocery store in Minnesota, he became fascinated with chemistry while musing on the variety of substances in the ingredients of the goods on the shelves. After a BS in chemistry at the Michigan College of Science and Technology, a PhD at the University of Minnesota and a postdoctoral period in Manchester, in 1937 he found himself in the chemistry department at the University of California in Berkeley. He was never to leave.

Several significant events in Berkeley were to have a profound effect on Calvin's future activities. It was in the Radiation Laboratory, which was founded in 1931 by Ernest Lawrence, that Sam Ruben and Martin Kamen discovered ^{14}C in 1940. Together with Andy Benson, they began to use the minute quantities then available to study the path of carbon in photosynthesis. The war put a stop to photosynthesis research; by the time it was over, Ruben had died in a laboratory accident, Kamen had left Berkeley for political reasons, and the availability of ^{14}C had grown enormously as the result of its production in nuclear reactors. Late in 1945, Lawrence, who then in effect controlled access to ^{14}C , suggested to Calvin that he make use of it, both to synthesize labelled compounds for medical research and to resume the photosynthesis exploration. In the heady post-war days of 'big' science in the style of the Manhattan project (in which the University of California had been involved), Lawrence made available to Calvin the funding and laboratory space that he needed.

Calvin never looked back. He was a man remarkable for a comprehensive knowledge of chemistry, unfailing originality and an ability to integrate information from a range of sources. Equally important was his talent for picking the right collaborators. The photosynthesis group was established in the Old Radiation Laboratory (ORL), an old wooden building that is affectionately remembered by the many graduate students and postdoctoral associates who worked there before its demolition in 1959.

Together with Benson, who was back in Berkeley after the war, Calvin set out to map the path of carbon.

In the basic experiments, $^{14}\text{CO}_2$ was fed in short bursts to illuminated green plants which were then studied for the progressive incorporation of radioactivity into product molecules. Initially, products were analysed by ion-exchange chromatography, but in 1949 the group was introduced to paper chromatography and progress then became very rapid. By the end of the 1950s — a few minor uncertainties aside — the carbon pathway lay revealed. The superb quality of the work was recognized in 1961 by the award to Calvin of the Nobel prize for chemistry.

Following the redevelopment of the ORL site, Calvin looked for funds to build a new laboratory. The former residents of ORL had no doubt that it must be a modern version of ORL, embodying its open and intimate quality. The nerve centre of ORL had been a large, white, Formica-topped table on which chromatograms and autoradiograms were laid out for contemplation; it is shown in the picture here. Despite the strong social cohesion of the group, with frequent parties and skiing or walking trips in the Sierras, the conversation round the big white table was almost always about research. (Calvin worried about those ski parties: would the guys be back in the lab on Monday morning?)

The new, three-storey, circular building was finished in 1963. Half of each of the upper two floors was a large semi-circular open laboratory. The big white table was installed at the centre of the third floor where it resumed its function as the group's meeting point; it was in exactly the same place when I last saw it in July 1996.

Resolution of the central biochemical question of photosynthesis, together with the space and facilities of the new building, opened new doors for Calvin. Retaining his earlier interests in chelates and other problems in physical organic chemistry, he

branched out into the biochemistry of learning, for a short while using the flatworm *Planaria*. After they had been taught to solve simple maze problems, it was claimed the animals could be chopped in half, whereupon the head portion grew a new tail, the tail end a new head with both reconstructed animals remembering their erstwhile 'parent's' training. Feeding trained *Planaria* to their naive cannibalistic relatives was said to confer new skills. (There was talk of feeding minced professors to the students!)

Deeply involved with origin-of-life questions even before probes were sent to the Moon and Mars, there seemed to be no limit to the breadth of Calvin's scientific curiosity. Early in the 1970s he established a group to work on cancer, and even after he relinquished the directorship of the group his ideas kept flowing. In the 1980s he seriously considered the use of certain plant oils as substitute diesel fuels should the world experience a severe petroleum shortage.

Calvin was overwhelmingly an experimental scientist. A consultant to chemical companies, until his Nobel prize and subsequent appointment to President Kennedy's Science Advisory Committee he had little interest in, or familiarity with, public affairs and politics. His many subsequent visits to Washington on public business broadened his understanding, yet he never thought of himself as, or wished to be, anything other than a scientist.

People sometimes talk of the great men and the great scientists they have known. For me and the many colleagues who worked with him, Melvin was *the* great man and *the* great scientist in our lives. Never hesitating to criticize in a scientific context, on a personal level he was always warm and caring. Those who knew him will never forget.

Vivian Moses

Vivian Moses is in the Division of Life Sciences, King's College, London W8 7AH, UK.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Common mechanism of infection by lentiviruses

Chemokine receptors are cofactors for infection with HIV-1 (refs 1, 2). HIV-2 and simian immunodeficiency virus (SIV) can infect feline cells expressing human CD4 (ref. 3), suggesting that feline cell-surface molecules can act as co-receptors for HIV infection. Here we demonstrate that the chemokine receptor CXCR4 mediates cell-cell fusion by cell-culture-adapted strains of feline immunodeficiency virus (FIV), indicating that there might be a common mechanism of infection by the primate and feline lentiviruses.

Cell-culture-adapted strains of FIV can infect and fuse with human cells⁴. Moreover, the ability of the FIV envelope glycoprotein (Env) to fuse with human cells is conveyed by a glutamate-to-lysine substitution in the V3 variable loop of its amino-acid sequence⁵, mirroring changes observed in the HIV V3 loop with the switch from a 'non-syncytium-inducing' to a 'syncytium-inducing' phenotype⁶, that is, to a phenotype that induces multinucleated cells. As the V3 loop is a primary determinant of the cell tropism of HIV and FIV, and may determine the nature of the chemokine receptor used by HIV (ref. 7), these data implicate feline chemokine receptors as candidate cellular receptors for FIV.

In the absence of reagents that can recognize feline chemokine receptors, we investigated fusion between persistently infected feline cells (CrFK) and human cells (HeLa). Mixing of FIV-infected CrFK cells with HeLa cells results in cell-cell fusion within 18 to 24 hours. We then screened monoclonal antibodies against cell-surface antigens for their ability to inhibit fusion. Incubation with the anti-human CXCR4 antibody 12G5 (ref. 8) significantly inhibits fusion at concentrations between 1 and 10 $\mu\text{g ml}^{-1}$, with maximal inhibition (>90%)

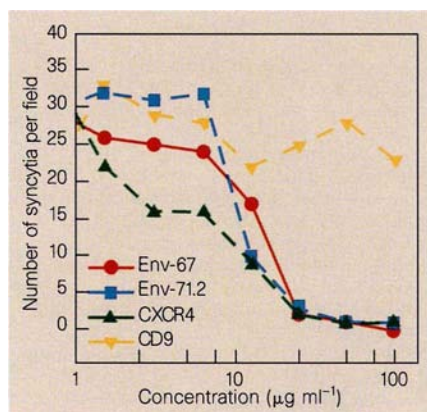


Figure 1 Effect of antibodies on FIV-mediated fusion of HeLa cells. HeLa cells were incubated for 1 h with either anti-feline CD9 (vpg15), anti-human CXCR4 (12G5), anti-FIV Env (Env-67) or anti-FIV Env (Env-71.2). FIV-producing CrFK cells were then added and incubated for 24 h. The number of syncytia per field was scored using a $\times 12.5$ periplan eyepiece (6.5 $\times$ 9 graticule). Syncytia were scored as multinucleated cells containing five or more nuclei per cell.

at concentrations from 25 to 100 $\mu\text{g ml}^{-1}$ (Fig. 1). Fusion is inhibited to a similar degree by the anti-FIV Env antibodies Env-67 and Env-71.2, confirming that the fusion is indeed FIV Env-specific. Incubation of the cells in the presence of a range of antibodies recognizing cell-surface antigens on either CrFK or HeLa cells does not affect fusion, even when they are included at concentrations equivalent to, or greater than, 12G5.

Similarly, fusion is not inhibited by incubation with a high-titre, neutralizing goat serum against the feline endogenous retrovirus RD114, but is inhibited efficiently by a polyclonal serum from an FIV-infected specific-pathogen-free cat. As the 12G5

antibody does not recognize feline cells expressing the feline homologue of CXCR4 (data not shown), these data indicate that the inhibitory activity might be mediated by binding of the 12G5 antibody to human CXCR4 on the surface of the human cells. Moreover, susceptibility to FIV Env-dependent fusion can be reconstituted by transfection of CXCR4-negative human cells with CXCR4 (B. J. W., L. Picard, M. J. H., J. D. T. & P. R. Clapham, manuscript in preparation), confirming that CXCR4 is indeed a cofactor for FIV Env-dependent fusion.

Our findings suggest that cell-culture-adapted strains of FIV use the CXCR4 chemokine receptor CXCR4 to infect and fuse with human cells. We believe that this is the first evidence of shared chemokine receptor use between primate and non-primate lentiviruses. FIV has puzzled AIDS researchers as it causes immunodeficiency with CD4⁺ T-cell depletion but has a broader host range which includes CD4⁻ cell types. It seems that the chemokine receptor may be a common link between FIV and the primate lentiviruses and a key determinant of the pathogenesis of AIDS.

Brian J. Willett

Margaret J. Hosie

James C. Neil

Department of Veterinary Pathology,
University of Glasgow,
Glasgow G61 1QH, UK

e-mail: b.willett@vet.gla.ac.uk

Julie D. Turner

James A. Hoxie

Rm 664, Clinical Research Building,
University of Pennsylvania,
415 Curie Boulevard,
Philadelphia, Pennsylvania 19104, USA

1. Feng, Y., Broder, C. C., Kennedy, P. E. & Berger, E. A. *Science* **272**, 872–877 (1996).
2. Dragic, T. et al. *Nature* **381**, 667–673 (1996).
3. McKnight, A., Clapham, P. R. & Weiss, R. A. *Virology* **201**, 8–18 (1994).
4. Pancino, G., Castelot, S. & Sonigo, P. *Virology* **206**, 796–806 (1995).
5. Verschoor, E. J. *J. Virol.* **69**, 4752–4757 (1995).
6. de Jong, J. J., de Ronde, A., Keulen, W., Tersmette, M. & Goudsmit, J. *J. Virol.* **66**, 6777–6780 (1992).
7. Choe, H. et al. *Cell* **85**, 1135–1148 (1996).
8. Endres, M. J. et al. *Cell* **87**, 745–756 (1996).

Volcanic iron, CO₂, ocean productivity and climate

The results of the IronEx II experiment show that, in the equatorial Pacific Ocean at least, the growth of unicellular marine diatoms can be dramatically increased by the addition of nanomolar quantities of inorganic iron to surface water^{1–4}. In common with the Southern Ocean and, to a lesser extent, parts of the northeast Pacific, these waters are termed 'high-nutrient, low-chlorophyll', meaning that the normal

nutrients nitrate and phosphate are found in significant quantities at the surface, unused by plankton. Addition of iron leads to uptake of much of these unused nutrients and an associated amount of inorganic carbon. Here I draw attention to further implications of the IronEx results. First, oceanic pH may increase in response to increased iron supply on a timescale of tens of thousands of years. Second, increased iron flux due to volcanic dust should increase oceanic productivity and hence atmospheric oxygen.

Though their application to the Southern Ocean remains untested, the IronEx

results add credence to the suggestion that the ecosystem of the Southern Ocean is iron-limited^{5,6}. They also indicate that an increase in the iron supply to that region, such as seems to have occurred during the last glacial period^{7,8}, would specifically benefit diatoms there. A general increase in productivity would rapidly (response time of roughly 100 years) tend to decrease atmospheric carbon dioxide^{9,10} because of decreased fugacity of CO₂ at the ocean surface. However, the stimulation specifically of diatoms (which fix organic but not calcite carbon) would further reduce atmospheric CO₂ by means of a change in the inorganic

to organic carbon 'rain ratio'¹¹ in detritus falling from the surface towards the sediment in all areas where iron has a limiting role.

The effect on atmospheric CO₂ of changes in the rain ratio was first described in ref. 11. An increased proportion of organic carbon reaching the sea bed from the surface results in a sediment that is more corrosive to calcite because, as the organic carbon is remineralized, it reduces the pH of the surface sediment. Increased dissolution of calcite leads, on a timescale of a few thousand years, to higher oceanic pH and hence lower atmospheric CO₂ concentrations. Support for this mechanism as a causal factor for lower glacial atmospheric CO₂ has recently come from proxy evidence for higher pH in the glacial ocean¹². No generally agreed reason has been advanced to explain why the rain ratio might change in this way on a global basis between glacials and interglacials, but the IronEx II result suggests that increased availability of iron in wind-blown dust during glacial time would have such an effect by stimulating diatom growth.

The eruption of Mount Pinatubo in June 1991 was a natural perturbation that allows a possible test of the hypothesis that ocean productivity is sensitive to atmospheric iron supply¹³. Fortunately, elegant observations¹⁴ of changes in the hemispheric gradient of O₂/N₂ ratio give some insight into patterns of global marine productivity during the period after the eruption. These observations are consistent with an extra pulse of the order of 10¹⁴ mol O₂ emanating from the Southern Hemisphere oceans, during austral summer 1991/92 (ref. 14).

The eruption injected crustal material (typically about 3 per cent iron by weight) into the troposphere and lower stratosphere and spread it over the globe. Most of the iron would have been carried on larger particles and would have been rapidly removed from the atmosphere, but some may have been carried far enough to enhance productivity in distant high-nutrient, low-chlorophyll regions, by far the largest of which is the Southern Ocean.

By November 1991, the Pinatubo stratospheric aerosol plume had reached those latitudes, a recent estimate of the mass deposition flux there being 9×10^{-13} g cm⁻² s⁻¹ at that time¹⁵. If 1 per cent of this flux was iron, sustained for 3 months over the area of the Southern Ocean, this would amount to roughly 4×10^{10} g iron. Given a typical carbon/iron molar ratio of 10⁵ for phytoplankton in iron-limited regions¹⁶, this would enable additional new production of about 7×10^{13} mol carbon. Such an increase would then give rise to the observed pulse of the order of 10¹⁴ mol oxygen into the atmosphere. What we know of the response of the ocean subsequent to the Pinatubo eruption is therefore entirely consistent with the notion that

'new production' in the Southern Ocean is sensitive to atmospheric iron input.

Andrew J. Watson

*School of Environmental Sciences,
University of East Anglia, Norwich NR4 7TJ, UK*

1. Coale, K. H. *et al.* *Nature* **383**, 495–501 (1996).
2. Behrenfeld, M. J. *et al.* *Nature* **383**, 508–511 (1996).
3. Cooper, D. J., Watson, A. J. & Nightingale, P. D. *Nature* **383**, 511–513 (1996).
4. Turner, S. M. *et al.* *Nature* **383**, 513–517 (1996).
5. Hart, T. J. *Discov. Rep.* **8**, 1–268 (1934).
6. Martin, J. H., Fitzwater, S. E. & Gordon, R. M. *Glob. Biogeochem. Cycles* **4**, 5–12 (1990).
7. Kumar, N. *et al.* *Nature* **378**, 675–680 (1995).
8. Martin, J. H. *Paleoceanography* **5**, 1–13 (1990).
9. Sarmiento, J. L. & Toggweiler, J. R. *Nature* **308**, 621–624 (1984).
10. Peng, T. H. & Broecker, W. S. *Nature* **349**, 227–229 (1991).
11. Archer, D. & Maier-Reimer, E. *Nature* **367**, 260–263 (1994).
12. Sanyal, A. *et al.* *Nature* **373**, 234–236 (1995).
13. Sarmiento, J. L. *Nature* **365**, 697–698 (1993).
14. Keeling, R. F., Piper, S. C. & Heimann, M. *Nature* **381**, 218–221 (1996).
15. Saxena, V. K., Anderson, J. & Lin, N.-H. *J. Geophys. Res.* **100**(D8), 16735–16752 (1995).
16. Sunda, W. G. & Huntsman, S. A. *Mar. Chem.* **50**, 189–206 (1995).

Hox genes misled by local environments

Cowing and Kenyon¹ raise an important issue concerning *Hox* genes by suggesting that these genes may not only be expressed according to position² but also according to cell lineage. This suggestion is based on two experiments in which the so-called M cell and two hypodermal V6 cells were manipulated to express a *Hox* gene, *mab-5*, outside the posterior region of the *Caenorhabditis elegans* embryo. Here we propose a less radical interpretation of these data.

In Cowing and Kenyon's first experiment, two blastomeres (EMS, P2) were laser-ablated in early embryos so that, owing to lineage transformation, eight rather than two V6 cells were produced. Three of the newly induced V6 cells expressed *mab-5* in anterior positions within the embryo, leading the authors to conclude¹ that the *Hox* gene is expressed independently of position and therefore according to lineage, in contrast to previous speculation³.

An alternative explanation for their result is that the newly induced V6 cells, which expressed *mab-5* despite being in an anterior position, were in fact in a posterior environment. *C. elegans* embryos are highly successful at maintaining their 'normal' body plan, even in partial embryos in which large numbers of cells have been transformed and whose morphology has been severely disturbed by ablated blastomeres^{4,5} (Fig. 2d of ref. 1). Only very extensive lineage transformation (of posterior into anterior) should hinder this process. Cowing and Kenyon¹ implied in their Figs 3a and 4 that this transformation always occurs after laser ablation of EMS and P2 blastomeres, on the basis of previous

studies^{4,6,7}. In fact, that work showed that such ablations only cause left–right transformations rather than large scale posterior to anterior (ABp to ABa) changes. Indeed, the normal mitosis of many cells that should have been transformed (Fig. 4 of ref. 1) indicates very incomplete transformations.

In the experiment, 11 of 44 expected new V6 cells were scored, of which only seven expressed *mab-5*. Three of these were positioned anterior to the region where the gene is normally expressed. It is conceivable that these transformed cells may have been accommodated, owing to incomplete transformations, in an environment supplying positional information, as shown in Fig. 1. So position may still be an important determinant of expression. The four V6 cells that did not express *mab-5* may show that position is important for *Hox* expression, as might the example where the wrong V4 cell, located posterior to V6, expressed *mab-5* (Fig. 4 of ref. 1).

In a second experiment, Cowing and Kenyon showed that the M cell can express *mab-5* outside the posterior of the embryo when its posterior migration is blocked. Analysing the migration of the M cell with a four-dimensional microscope⁸ we learned that it originates at 35% egg length (roughly 3.5 h) and migrates to 40% in the next half-hour. Cowing and Kenyon¹ report that 7% of control embryos express *mab-5* at 4 h of development. The gene has already been expressed normally at 40% egg length, so interference with cell migration did not, in

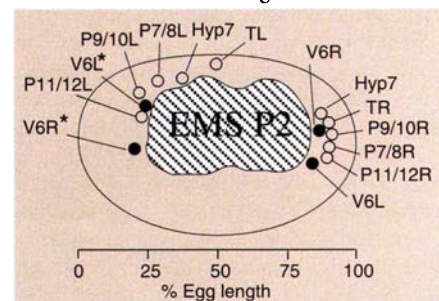


Figure 1 Positions of posterior hypodermal cells¹⁰ in an EMS- and P2-ablated embryo after 5 h of development. Extensive lineage analysis showed that this embryo, although P2-ablated, expressed the normal ABp-derived cell fates and so four V6 cells are present⁷, not eight as expected by Cowing and Kenyon¹. Two posterior regions exist in this embryo. The normally posterior cells of the left side are trapped anterior to the ablated blastomeres. The ectopically induced V6 cells (asterisks) are in an anterior position above cells that normally demarcate a posterior position in the embryo. The T, Hyp7 and P cells in this embryo are derived from untransformed ABp and ABpr lineages which were also normal in the embryo reported by Cowing and Kenyon (clone 1 in their Fig. 4)¹, as only EMS was ablated. The fates of these lineages was not investigated in the second embryo (clone 6) with anterior expression of *mab-5* (ref. 1).

fact, shift the expression of *mab-5* to the anterior of the embryo.

As both experiments of Cowing and Kenyon are open to alternative interpretations it seems premature to change our thinking about the expression of *Hox* genes⁹.

Ralf Schnabel, Heinke Schnabel

Max-Planck-Institut für Biochemie,
Am Klopferspitz 18a,
82152 Martinsried, Germany

1. Cowing, D. & Kenyon, C. *Nature* **382**, 353–356 (1996).
2. Lawrence, P. & Morata, G. *Cell* **78**, 181–190 (1994).
3. Kenyon, C. *Cell* **46**, 477–487 (1986).
4. Hutter, H. & Schnabel, R. *Development* **120**, 2051–2064 (1994).
5. Priess, J. R., Schnabel, H. & Schnabel, R. *Cell* **51**, 601–611 (1987).
6. Moskowitz, I. P. G., Gendreau, S. B. & Rothman, J. H. *Development* **120**, 3325–3338 (1994).
7. Hutter, H. & Schnabel, R. *Development* **121**, 1559–1568 (1995).
8. Hird, S. & White, J. G. *J. Cell Biol.* **121**, 1343–1355 (1993).
9. Emmons, S. *Nature* **382**, 301–302 (1996).
10. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. *Dev. Biol.* **100**, 64–119 (1983).

The fossil *Ainiktozoon* is an arthropod

The enigmatic fossil *Ainiktozoon loganense* Scourfield 1937 (ref. 1) from the Lower Silurian period of Lesmahagow, Scotland, has puzzled most who have studied it. We have examined several new specimens, some exceptionally well preserved and with soft tissues, from the type locality, and conclude that this fossil is an arthropod, not a protochordate, with affinities to the Thylacocephala.

Discovered in the late nineteenth century, *Ainiktozoon* was not described for nearly 50 years². When Scourfield first published a detailed description¹, he concluded that it was probably allied with the Urochordata. Subsequent workers considered *Ainiktozoon* only with regard to the early history and origin of the chordates³. A chief difficulty in interpreting *Ainiktozoon* has been in the apparent lack of homology of its known features with those in other organisms. We report here on new specimens from the type locality in which three-dimensional soft tissues are preserved, including organic (probably cuticular) structures and phosphatized musculature.

Ainiktozoon fossils are usually found as laterally flattened compressions on bedding planes of compacted laminated siltstones throughout the Jamoytius Horizon in the Patrick Burn Formation. Specimens, of which RMS 1996.31.1 (Fig. 1a) and RMS 1977.3.7 (Royal Museum of Scotland) are among the best-preserved examples, range in length from 25 to 240 mm. Scourfield originally drew the fossil with a fringe in a dorsal position, on what he termed a sac-like body¹, an attitude followed by Ritchie² (Figs 2 and 6 of ref. 2). Only after we turned these reconstructions upside down did we

realize that *Ainiktozoon* could have arthropodan affinities.

In our new interpretation (Fig. 1b), these animals display an anterior carapace covering a segmented body which bears eight biramous, paddle-like limbs and short, terminal, caudal rami at its posterior end. These trunk segments are marked by longitudinal muscle fibres interrupted with segmental, connective tendons. Originating from these tendons are diagonally arranged limb muscles that insert into the paddle-like limbs. The 'capsule' of Ritchie², which lies anterior to the segmented region, emerges now as a highly muscular foregut. We reinterpret the various 'fascicles' noted by previous workers as phosphatized carapace adductor muscles, dilator stomach muscles and other soft anatomy associated with an arthropod cephalon.

The carapace apparently was rather thin in life and is usually only poorly preserved.

In two specimens (RMS 1996.31.1 and RMS 1977.3.7) the carapace seems not to have the distinct bivalved form of other thylacocephalans but rather is splayed out with left and right valves revealed. Both of these specimens show pairs of compound eyes with hexagonal facets; typical arthropodan features. The eyes are among the smallest of any thylacocephalan but are otherwise similar to those found in other members of this group, as yet undescribed, from the Silurian of Waukeschau, Wisconsin⁴.

Specimen RMS 1996.31.1 also preserves anteriorly a set of 'jack-knifed', superimposed, subchelate, raptorial limbs. The segments are rather large and armed with robust spines, and clearly resemble the anterior raptorial, subchelate limbs of the Thylacocephala. They are similar in form to limbs on the American Silurian material (mentioned above) as well as those seen on

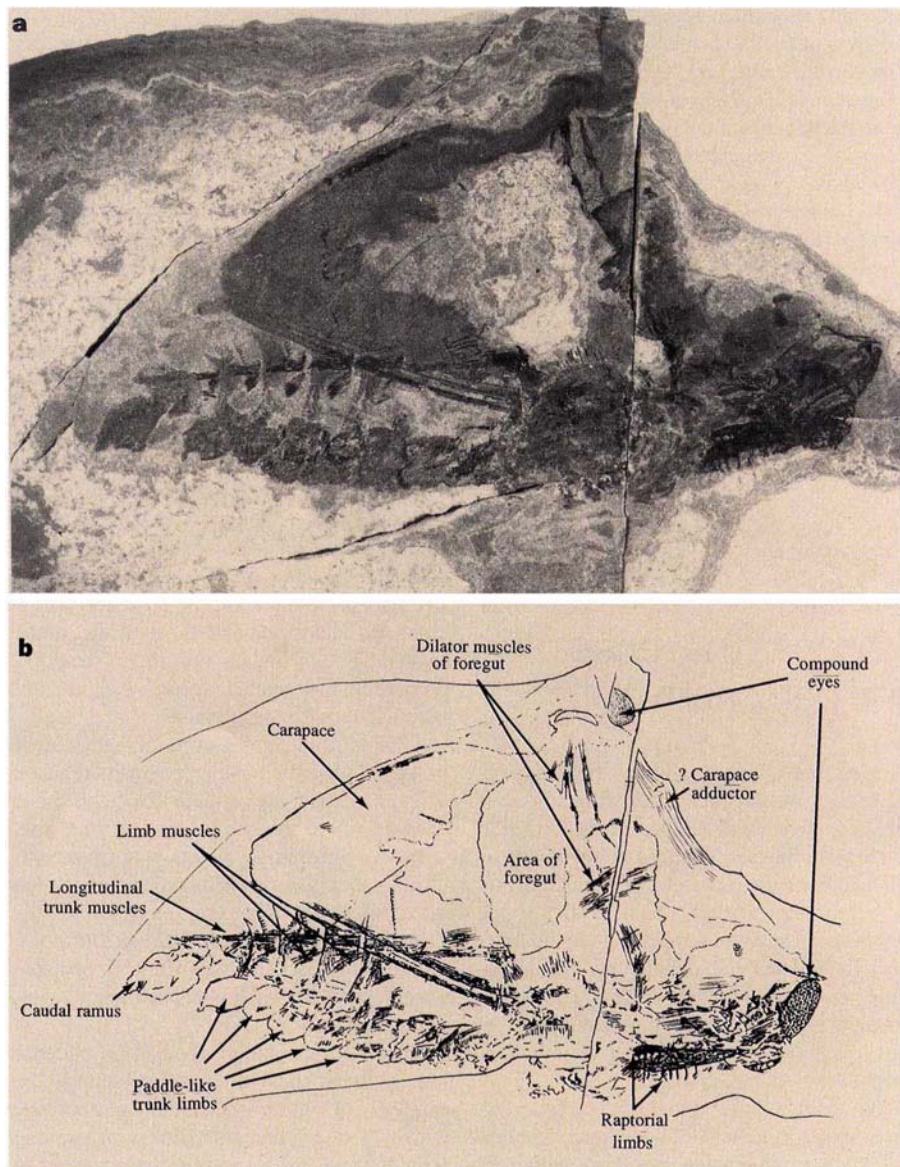


Figure 1 a, New specimen of *A. loganense*, immersed in water, with preservation of non-mineralized tissues, including cuticle and musculature. Specimen RMS 1996.31.1. b, Camera lucida drawing of RMS 1996.31.1, with thylacocephalan arthropod features identified. The length of the specimen from compound eye to caudal ramus is roughly 18 cm.

the Jurassic thylacocephalan *Dollocaris*⁵. Beneath the carapace on RMS 1977.3.7, pairs of gills with the same form and position of those of *Dollocaris* are visible.

All the pertinent features^{6,7} of the Thylacocephala are present: large carapace, paired compound eyes, subchelate limbs on the anterior part of the body, large laminar gills beneath the carapace, and a segmented, posterior trunk with swimming limbs. If we have correctly interpreted these structures, then *Ainiktozoon* certainly resembles a thylacocephalan arthropod more than it does any known protochordate, a relationship previously only hinted at by Dzik⁸.

Ainiktozoon also possesses several unique features (for thylacocephalans) related to the apparently flattened form of the carapace and the relatively small size of the eyes, and so may represent a new family within that group. All other features of these new specimens accord well with what is known of the Thylacocephala.

Wim van der Bruggen

26 Garven Road, Stevenston KA20 3NX, UK

Frederick R. Schram

Universiteit van Amsterdam,

Instituut voor Systematiek en Populatiebiologie,

1090 GT Amsterdam, The Netherlands

David M. Martill

Department of Geology, University of Portsmouth,

Portsmouth PO1 3QL, UK

e-mail: martilld@geol.port.ac.uk

1. Scourfield, D. J. *Proc. R. Soc. Lond. B* 121, 533–547 (1937).
2. Ritchie, A. *Alcheringa* 9, 117–142 (1985).
3. Tarlo, L. B. H. *21st Int. Geol. Congr. Scand.* 22, 113–123 (1960).
4. Mikulic, D. G., Briggs, D. E. G. & Kluessendorf, J. *Phil. Trans. R. Soc. Lond. B* 311, 75–85 (1985).
5. Secretan, S. & Riou, B. *Ann. Paleontol.* 69, 59–97 (1983).
6. Rolfe, W. D. I. *Trans. R. Soc. Edinb.* 76, 391–399 (1985).
7. Rolfe, W. D. I. *Acta Zoologica (Stockh.)* 73, 301–304 (1992).
8. Dzik, J. *Dzieje Zycia Na Ziemi. Wprowadzenie Do Paleobiologii* 1–464 (PWN, Warsaw, 1992).

D2 dopamine receptors and personality traits

Temperament, a person's basic behavioural patterns or personality traits, differs greatly between individuals. Experimental studies of animals have implicated the dopamine neurotransmission system in the initiation of behaviour, reward and motivational processes¹. Here we report that in normal human subjects the density of D2 dopamine receptors in the brain correlates strongly with a detached personality — a trait that includes lack of closeness and warmth in personal relations.

We studied the brains of 24 normal Caucasian adults (aged 18–38 years, 14 men and 10 women) using positron emission tomography (PET). We determined the D2-dopamine-receptor density in the putamen, one of the basal ganglia, using a Scatchard procedure based on two PET measurements

with the radioligand [¹¹C]raclopride². The extent of [¹¹C]raclopride binding in the left and right putamen was averaged, and the most ventral part of the putamen was excluded to minimize a possible contribution from D3-dopamine-receptor binding in the ventral striatum. We measured personality traits of the subjects using the 'Karolinska scales of personality' (KSP), a self-report questionnaire³.

The D2-receptor density varied between 18 and 37 pmol ml⁻¹ (29.1±6.9; mean±s.d.). The density strongly correlated with the Karolinska scale of 'detachment' (one of 15 personality scales) with $r = -0.68$, $P < 0.001$ (Fig. 1), a correlation which remained significant even after a Bonferroni correction for 15 comparisons, and irritability ($r = -0.51$, $P < 0.01$). The 'detachment' scale includes the tendency to avoid giving and taking confidences and to avoid involvement with other people.

In a recent study⁴, using the 'inventory for interpersonal problems', individuals who scored high on the detachment scale described themselves as cold, socially aloof and vindictive in their relationships, whereas those who scored low reported problems with being overly nurturing and exploitable. Thus both extremes of this scale describe individuals who may have considerable interpersonal problems in daily life.

Detachment can encompass social isolation, indifference to other people and lack of intimate friendships, traits included among the category of 'negative' symptoms that commonly characterize patients with schizophrenia⁵; a terminology derived from the assumption that negative symptoms represent a loss of function. Classical antipsychotic drugs block D2 dopamine receptors⁶ and are effective for the treatment of positive schizophrenic symptoms such as hallucinations, delusions and thought disorders. In contrast, negative symptoms respond poorly, or may even worsen, when treated with D2-receptor antagonists⁷.

PET studies on D2 dopamine receptors in schizophrenia have previously focused on neuroleptic-naïve patients, who have been admitted to hospital for the first time, with predominantly positive symptoms^{8,9}. Our finding of a strong association between detachment and D2-receptor density indicates that it may prove fruitful to look for a low density of D2 receptors in schizophrenic patients with predominantly negative symptoms.

Behavioural genetic studies on rodents indicate that individual differences in the number of dopamine-releasing neurons, within a dopamine cell group, are maintained across all dopamine cell groups in the brain (for review see ref. 10). This is reflected in a strong correlation between the reactivity of hormonal and behavioural indicators in different dopamine cell

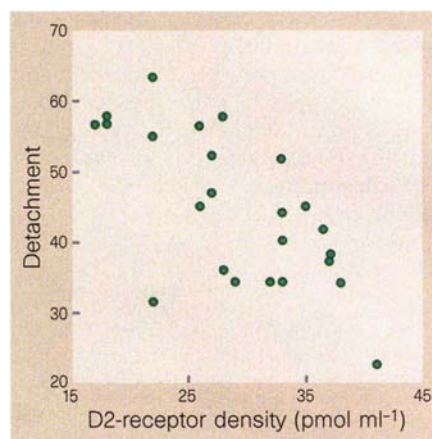


Figure 1 Individual values ($n = 24$) for D2-dopamine-receptor density plotted against KSP detachment scores. To adjust for the effect of gender, the scores were transformed to T scores using normative data. The T scores have a mean ($\pm$ s.e.m.) of 50 (± 10) in the normal population.

groups. We assume that the individual differences in D2-receptor density that we have found in the putamen are maintained in mesolimbic and neocortical regions.

Each of the five dopamine receptor subtypes (D1–D5) has a distinct pharmacological profile and a unique neuroanatomical distribution¹¹. They may therefore be biological markers for different functional aspects of dopamine activity, and attempts have been made to associate polymorphic variants of dopamine receptor subtypes with features of human personality^{12–14}.

As yet it is unclear whether individual variability in D2-dopamine-receptor density is wholly genetically determined or whether it is subject to environmental influences. We propose that neuroreceptor density is a useful biochemical measure for relating the genetic endowment to human personality traits. This approach provides a new way to search for empirical biological correlates of personality in man.

Lars Farde, J. Petter Gustavsson

Erik Jönsson

Department of Clinical Neuroscience,

Karolinska Hospital, S-171 76 Stockholm, Sweden

e-mail: larsf@psyk.ks.se

1. Mirenowicz, J. & Schultz, W. *Nature* 379, 449–451 (1996).
2. Farde, L., Hall, H., Pauli, S. & Halldin, C. *Synapse* 20, 200–208 (1995).
3. Schalling, D., Åsberg, M., Edman, G. & Orelund, L. *Acta Psychiatr. Scand.* 76, 172–182 (1987).
4. Weinryb, R. M. *et al. Pers. Individ. Diff.* 20, 13–23 (1996).
5. Andreasen, N. & Olson, S. *Arch. Gen. Psychiatry* 39, 789–794 (1982).
6. Farde, L., Hall, H., Ehrin, E. & Sedvall, G. *Science* 231, 258–261 (1986).
7. Meltzer, H., Sommers, A. & Luchins, D. J. *Clin. Psychopharmacol.* 6, 329–338 (1986).
8. Wong, D. F. *et al. Science* 234, 1558–1563 (1986).
9. Farde, L. *et al. Arch. Gen. Psychiatry* 47, 213–219 (1990).
10. Depue, R., Luciana, M., Arbsi, P., Collins, P. & Leon, A. J. *Pers. Soc. Psychol.* 67, 485–498 (1994).
11. Meador-Woodruff, J. *Ann. Clin. Psychiatry* 6, 79–90 (1994).
12. Ebstein, R. *et al. Nature Genet.* 12, 78–80 (1996).
13. Benjamin, J. *et al. Nature Genet.* 12, 81–84 (1996).
14. Malhotra, A. *et al. Mol. Psychiatry* 1, 388–391 (1996).

The campaign for biodiversity

The Idea of Biodiversity: Philosophies of Paradise

by David Takacs

Johns Hopkins University Press: 1996.

Pp. 393. \$35.95, £25

Paul Colinvaux

David Takacs writes of 'biodiversity' as seen by the newish discipline of 'science studies'. For him the word 'biodiversity' signals a cultural phenomenon, to be explored in the writings of prominent biologists and by interviews. As his title suggests, the result is an account of an idea rather than a discourse on variety.

We naturalists have a visceral desire to conserve as large an array of species as possible. We know that to do this we must conserve habitats as well as animals or plants, and we are absolutely sure that this diversity of life and landscape has deep meaning for human lives. We have had need of a catchphrase to describe the values we would preserve. 'Endangered species' can no longer carry whole ecosystems on their backs, as snail darters finally failed to do in Tennessee and spotted owls do so tenuously in the forests of the Pacific northwest. 'Wilderness' can have many meanings. So let 'biodiversity' carry the load. Conserve 'biodiversity' and you preserve species richness, habitats, and opportunities for living — all three.

Takacs dates the campaign for biodiversity to a 1985 forum of the US National Academy of Sciences when prominent biologists called a press conference to say: "The species extinction crisis is a threat to civilization second only to the threat of thermonuclear war". I incline to reverse the importance of those two threats, but the point is made. Now, what to do about it?

Here's the rub: people do not value biodiversity, despite the urging of scientist and humanist that it has intrinsic or aesthetic worth. We can argue, for instance, that useful chemistry can be borrowed from the prodigious variety of compounds used by plants and invertebrates in their evolutionary games of attack and defence, but the largest sum so far extracted from the pharmaceutical industry for drug hunting is a paltry \$1 million. Tourism perhaps offers greater bargaining power. In Costa Rica, for example, biodiversity now earns as many (or as few) export dollars through tourism as does the sale of bananas.

When Takacs asked ecologists (broadly defined) if the vast array of species has ecosystem value, we should not be surprised that he got the answer, "Yes". A result is that the theme of 'complexity begets stability' keeps reappearing in the book. The fear that loss of diversity has practical costs is at its

most frightening when we acknowledge that we do not know what most living things do, or even what they are. The argument from our ignorance is, "Do not throw out what you do not understand or you will be sorry". But Robert May and others have long since shown us that the mathematical and logical grounds for linking complexity to stability are illusions. Takacs' investigations suggest that we yield those illusions only with difficulty.

The book does not attempt to review diversity theory. There is no mention of the work of Robert MacArthur or G. E. Hutchinson, nor even of the giant debates between the followers of Clements and Gleason. This is a work of 'science studies', which seems a form of journalism: the awareness is now, the time-horizon is short, the output might be policy advice.

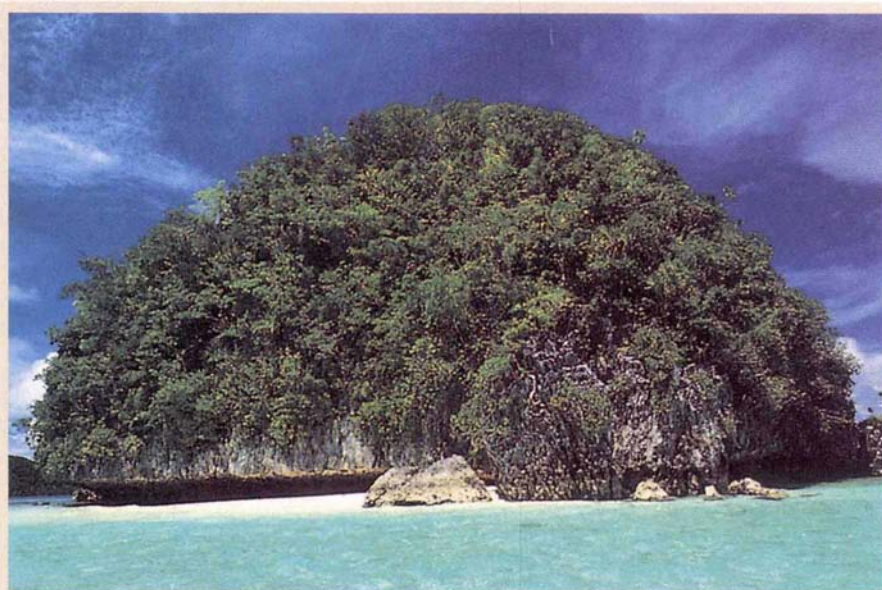
The plans and hopes of our most articulate advocates for biodiversity emerge from the book. Daniel Janzen has offered the most detailed plan for the tropics — we should sell rights to exploit plant chemistry and use the money to buy land and educate people. Following this plan, Costa Rica was able to negotiate that \$1 million from the Merck company. The country now has a programme to train ordinary people, from *campesinos* to clerks, to be parataxonomists for pay — that is, they collect insects in a dis-

criminating way for the national inventory. The real pay-off is the creation of a growing nucleus of people aware of and enthusiastic about biodiversity.

Janzen's growing army of parataxonomists is a practical use of what E. O. Wilson calls the latent 'biophilia' in people, an innate love of nature (read 'biodiversity') that can be tapped for support. Wilson's prominent role in the campaign for biodiversity is celebrated in a separate chapter describing his attitudes and work.

The journalistic methods of 'science studies' result in a 100-page documentary chapter called 'Values' which prints, apparently verbatim, the answers from Takacs' jury of a dozen or so biologists and conservation writers when asked provocative questions about motivation and belief. These are transcribed from tape recordings, broken syntax and all.

The questions in an interview determine the scope of the answers and I find little evidence of questions about people whose very beliefs are contrary to the ethics of those of us who want to conserve biodiversity. What of those whose creation myths cause them so to abhor a biologist's vision that they try to keep evolutionary learning out of schools? What of the ethic of asking tropical peoples to spare their forests when we have turned England into a safe arable garden, or when a



Islands of wilderness on a crowded planet

"Palau, where tropical forests grow in voluptuous proximity to coral reefs, is like no other place on earth," writes David G. Campbell. "It is a dispersed archipelago of 340 islands, volcanic nubbins on the western edge of Micronesia. There is plenty of wilderness

here; Palau's population is only 16,000." Ecologist Campbell travelled the globe in search of wilderness areas, and he provides vivid portraits of ten such places he found in *Islands in Space and Time* (Houghton Mifflin, \$35).

US Congress wants to cut down old-growth forests to save jobs? How far can we train parataxonomists when we cannot even persuade our societies to pay the salaries of taxonomists who would catalogue the diversity we are trying to save?

Perhaps the journalism of science studies has no answers to these things because we have no answers ourselves. □

Paul Colinvaux is at the Smithsonian Tropical Research Institute, Box 2072, Balboa, Republic of Panama.

Complex physics

Infinite Potential: The Life and Times of David Bohm

by F. David Peat

Addison-Wesley: 1996. Pp. 353. \$25

Chris Philippidis

David Bohm (1917–94) had an influence on science in general, and physics in particular, that could be thought of as subterranean. On the occasions when it broke through the surface, it did so in stunningly unexpected ways. The effect was always to change the scientific landscape. His three most significant technical interventions dealt with plasma physics, hidden variables in quantum mechanics and the role of electromagnetic potentials in quantum mechanics. His impact on the philosophy of science was profound, and several other disciplines felt the effects.

But, as David Peat's biography shows, Bohm's home ground was the foundations of physics and quantum theory. In this realm he respected no taboo nor the sanctity of any concept; he wrestled with ideas and notions that he teased out from the recesses of a variety of disciplines and he remoulded them into new meanings. Guided by his understanding of dialectical materialism and quantum theory, he unravelled a vision of an infinitely complex world, whose potential manifestations are inexhaustible in their diversity and therefore representable only through partial and provisional theorizations. Out of such primeval stuff Bohm hewed his physics. All his contributions were hallmarked by such radical notions as interconnectedness, active and passive information, implicate and explicate orders and so on.

Bohm was turned into an outsider and driven underground by the historical circumstances that shaped the greater part of this century. The Maxwells and Boltzmanns, the Plancks and Einsteins were in the business of laying down designs for theoretical edifices and providing specifications for their foundations. By the time the Second World War was over, however, this way of doing physics had already been undermined by the emergence of corporate scientific enterprise. The appearance of quantum theory in the 1920s created in its wake the 'Copenhagen group' and its

'board of directors' — Niels Bohr, Werner Heisenberg, Wolfgang Pauli and others.

The Copenhagen interpretation established itself effectively as a 'brand name' guaranteeing that all fundamental questions about the micro-world had been taken care of once and for all. The period after the war saw a generation of physicists who had been weaned on the belief that there were no more big questions to be addressed. For them, physics was a matter of cashing in on the rich pickings that came from the application of quantum theory. Despite his membership of the scientific elite at Princeton, Bohm's insistence on unpicking the meanings of the Copenhagen interpretation in the early 1950s looked anachronistic and defined him as an outsider. The reactionary political persecutions in the United States at the time aided his exclusion from the scientific establishment. These events set the pattern for the remainder of his life. His strong affinity with Einstein is symbolic of Bohm's subsequent place in physics.

If a biography can be considered to be a map of someone's path through life, then mapping out Bohm's must be a particularly daunting affair. What symbols and contours, for instance, could adequately image a thinker who insisted that images have a capacity to delude? How would a biographer fix the identity of a subject who warned against the illusion of capturing an essence? A purely scholarly biography of Bohm would be duty-bound to square up to such paradoxical and problematic issues. If, in addition, it sought to register some solidarity with his views, then it would also have to resolve the added complication of being critical of its own analysis.

Peat's narrative essentially sidesteps these intellectual hazards. The book is obviously aimed at a general readership, and is therefore forced to cover an enormously wide subject for a diverse audience. Bohm's contributions range over a spectrum of topics from physics to religion, while those interested in him come from all walks of life and have all sorts of expectations. Peat's professionalism in writing and experience with popular genres are well up to the task. His understanding of Bohm's work is extensive and full of insight and he is able to represent it in a compelling and readable way. He does not shy away from explaining the physics and the philosophy, but he inevitably cannot allow himself to dwell on either. The effect is sometimes reminiscent of 'doing Paris in the morning and Rome in the afternoon', and the specialist scientist may yearn for more detail.

Peat's personal knowledge of Bohm provides him with anecdotes to draw on. He is able to write experientially and recount the story with a conviction that would not be available to an author who knew the subject only at second hand. On the whole he uses this resource skilfully and sensitively to paint a vivid 'warts and all' portrait. On a few occasions, however, he invokes some banal inci-



Bohm: changed the scientific landscape.

dents that risk lowering the tone of his narrative. Luckily, these are sufficiently rare not to detract from its overall quality.

Over the years Bohm has been claimed by many followers and misrepresented by many interpreters, some of whom were quite close to him. Bohm himself kept different parts of his life separate, and thus contributed to this state of affairs. Peat's book is an important and welcome attempt to begin the process of reconstructing a complex personality that has hitherto appeared paradoxically fragmented and incomplete. □

Chris Philippidis is in the School of Interdisciplinary Sciences, University of the West of England, Bristol BS16 1QY, UK.

Seeing is believing

The Biological Universe: The Twentieth-Century Extraterrestrial Life Debate and the Limits of Science

by Steven J. Dick

Cambridge University Press: 1996. Pp. 578. £40, \$54.95

Stephen G. Brush

In 1982, Steven J. Dick published *Plurality of Worlds: The Origins of the Extraterrestrial Life Debate from Democritus to Kant*, covering two millennia of speculation in about 240 pages. Then, in 1986, Michael J. Crowe published *The Extraterrestrial Life Debate, 1750–1900: The Idea of a Plurality of Worlds from Kant to Lowell*, digesting a sesquicentury of philosophical, religious and scientific argument in 680 pages. Now we have the final volume of the Cambridge University Press trilogy, a detailed account of 100 years of research and controversy, from Percival Lowell's claimed but refuted discovery of life on Mars (in 1894) to the claimed and confirmed discovery of a planetary pulsar

system by A. Wolszczan and D. A. Frail (in 1992). Together, these three books provide, in a way that is accessible to readers with a modest scientific background, a fascinating and definitive account of terrestrial humans' efforts to answer the question "Are we alone in the Universe?"

After a brief recapitulation of the story from Democritus to Lowell, *The Biological Universe* starts with the anthropocentric thesis argued in 1903 by Alfred Russel Wallace in his book *Man's Place in the Universe*. Wallace's conclusion was that because a very complex combination of peculiar conditions was required to develop and maintain life on Earth, "it seems in the highest degree improbable" that these conditions "can all be found again combined" anywhere else, hence Earth is the only inhabited planet in the Universe. Dick later compares this argument with those used in formulating the anthropic principle in the 1970s and 1980s, suggesting that they share a teleological justification for the uniqueness of humanity's existence.

Dick shows how changing theories of the origin of our own Solar System affected assumptions about the number of planetary systems in the Universe, and thus indirectly about the prevalence of extraterrestrial life. For example, James Jeans argued in the 1920s and 1930s that because (according to the theory developed by Thomas Chamberlin and Forest Moulton in the United States, and independently by Jeans and Harold Jeffreys in England) planetary systems can

be formed only by the close encounter of two stars, a very rare event, we are probably alone in the Universe.

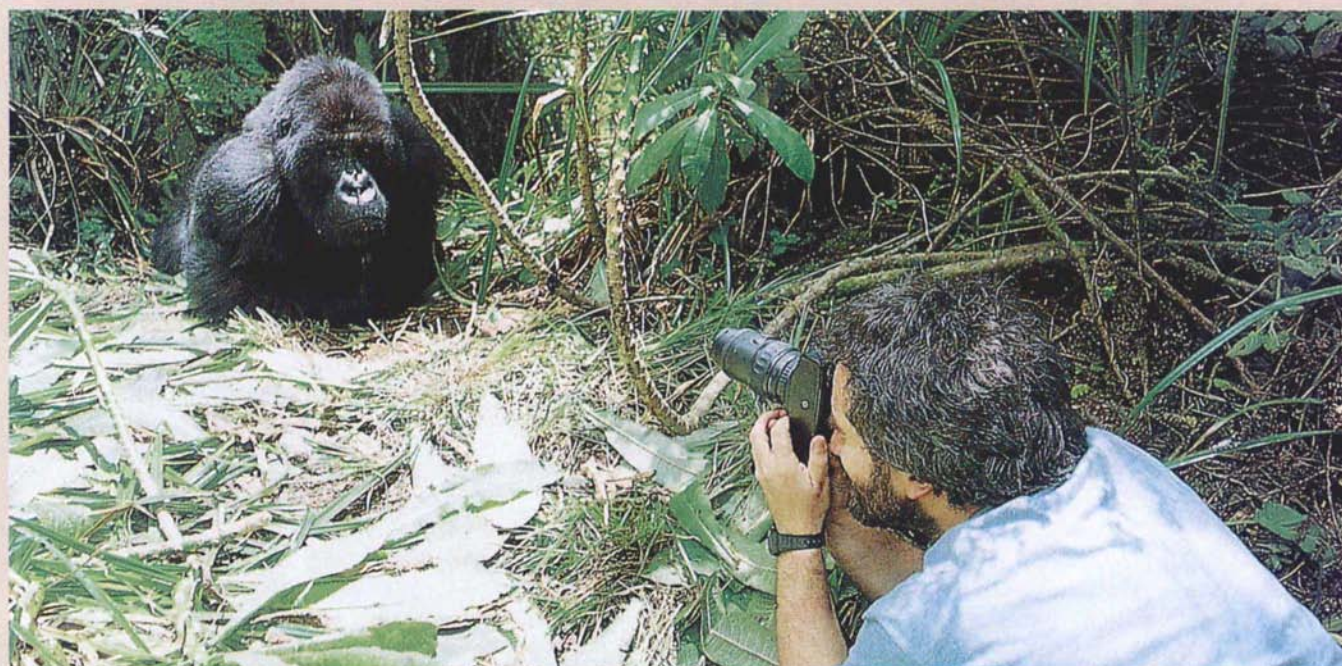
The revival of the nebular hypothesis after 1940 led many astronomers to believe (or perhaps, in a few cases, resulted from their wish to believe) that the formation of planets is a normal part of star formation, so there must be many habitable planets. This revival was due not only to the theoretical defects of the encounter theory and alleged discoveries of extrasolar planets in the 1940s, as suggested by Dick, but also to Cecilia Payne-Gaposchkin's discovery of the abundance of hydrogen in stars and Hannes Alfvén's proposal of momentum transfer from the Sun to planets by magnetic braking; these two ideas allowed theorists to correct the deficiencies of the original Laplace theory.

A major theme of the book, however, is the recognition of the limited value of theory and the increasing emphasis on observation. Decades of observations of Mars debunked Lowell's theories about intelligent life there. Decades of analysis of the proper motions of selected nearby stars generated and then deflated hopes that tiny variations in those motions could prove that the stars are accompanied by planets. (Both the 1995 discovery by Michel Mayor and Didier Queloz of a planet orbiting 51 Pegasi, confirmed by Geoffrey Marcy and R. Paul Butler, and the claimed but still disputed 1996 discovery of evidence for Martian life in a meteorite, were too late to be discussed in the text, but the former earned a brief

notice in the bibliographical essay at the end of the book.)

There is a full account of the Search for Extraterrestrial Intelligence (SETI) by monitoring radio signals from nearby stars, originally stimulated by the proposal of Giuseppe Cocconi and Philip Morrison. This was published in September 1959 in *Nature*, "which had a reputation for printing speculative papers". By 1992, SETI was respectable enough to become the focus of an observing programme sponsored by NASA, although the commitment of the US government to funding this programme proved to be quite capricious.

From a sociological viewpoint the most interesting topic is the controversy about unidentified flying objects (UFOs). A first wave of reports of mystery spaceships in 1896–97 elicited suggestions from the *St Louis Post-Dispatch* and other American newspapers that Martians had come to visit us. After a 50-year hiatus, many sightings of flying saucers occurred, beginning in 1947; there were more than a thousand reports in each of the years 1952, 1957 and 1966. (Was it only Americans who could see flying saucers, as the author seems to imply?) Significantly, Dick places his chapter on UFOs immediately after a chapter on extraterrestrials in literature and the arts. He discusses the reluctance of almost all scientists to offer any public comment on this alleged phenomenon; he ascribes their caution to fear of risking their careers by getting involved in a controversy that was increasingly domi-



Face to face with a close relative

"Mountain gorillas have become media superstars over the past 20 years," writes Michael Leach. "For many people, the ultimate wildlife experience would be to see the mountain gorilla."

Leach must have done his gorilla watching before the recent wars in Rwanda and Zaire. This picture is from his popular account of mankind's closest relatives, *The Great Apes: Our Face in Nature's Mirror* (Blandford, \$29.95, £25). The book looks at the animals' behaviour in the wild, treatment in captivity, and conservation status.

nated by hoaxes, fantastic claims and mysticism. Insofar as the idea of extraterrestrial life became associated with UFOs, scientists were discouraged from taking it seriously.

Modern readers may be surprised to learn that before the middle of this century "theories of the origin and evolution of life played very little role in the extraterrestrial life debate". Aside from a very few biologists like Wallace, most of the participants in that debate were astronomers (or, earlier, philosophers). Dick nevertheless looks at these theories as a background to recent research in cosmic evolution and the formation of a new discipline that has not quite decided whether to call itself 'exobiology' or 'bioastronomy' or some other hybrid term.

In his concluding chapters, Dick reflects on how the extraterrestrial life debate has affected, or is likely to affect, our ideas about the limits of science, the nature of evidence, God's role in the Universe and the meaning of life. Without accepting Dick's views on those subjects, every reader should be grateful for a book that succeeds in enriching our knowledge of the historical and factual background needed for any serious discussion of these 'big questions'. □

Stephen G. Brush is at the Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA.

The health of nations

The People's Health: A Memoir of Public Health and Its Evolution at Harvard

by Robin Marantz Henig

National Academy Press: 1996. Pp. 256.

\$29.95, £24.95

John Galloway

'Public health' tends to enjoy a similar bad press to its big brother, 'public interest'. Neither is viewed enthusiastically by the public whose well-being they claim to serve. Both are seen as notions dreamed up by politicians with shabby motives, usually excuses for infringing people's personal freedom.

In reality the public — as opposed to the private — approach to health has much to recommend it. Not only because it is rooted in the precept that prevention is better than cure — and certainly better than no cure. But, as the dean of the Harvard School of Public Health, Harvey Fineberg, perceptively says in *The People's Health*, "in my heart of hearts, I do really believe we will never in this world solve health problems one person at a time". The school's history gives continuity to this book which is written rather racily by the journalist Robin Henig.

Fineberg's remark is bold for someone living in the United States where the "one person at a time" approach to health is the

philosophy underpinning the country's biggest industry — an industry that denies its benefits to a significant number of the population. Equity of access to health-care is apparently too expensive for the government of the richest country in the world, either in dollars or, more significantly, in votes. People will vote for their own health, but not their country's if it includes the 'badly off'.

As Henig points out, the fundamental flaw of private insurance schemes as a means of guaranteeing health is that the more you are at risk, the less likely you are to get cover. Whatever its present difficulties, the reasoning behind the creation of the United Kingdom's National Health Service was sound and humane and removed the financial worries that used to accompany ill health.

The twentieth century has been called the people's century, the age of mass everything, including health. The book is a roll-call of infectious diseases, once national scourges, that have largely succumbed to affluence: better housing, hygiene, food, medicine, education — and, by way of a *coup de grâce*, mass vaccination and immunization. It is salutary to be reminded that in the mid-nineteenth century half the United Kingdom's population died before the age of 14.

With the benefit of hindsight, battles against infectious diseases can be seen as skirmishes in a continuing war. Syphilis has been replaced by AIDS, tuberculosis by lung cancer and malnutrition by heart disease. Infectious diseases are becoming resistant to drugs. Stranger and more refractory diseases have appeared than medicine is used to — AIDS and Creutzfeldt-Jakob disease. New viruses have emerged as overpopulation forces mankind to invade stable ecologies. But most invidious of all is the realization that our genes are now the enemies within and our national economies are the enemies without. Cancers are genetic diseases, but a third are triggered by diet and another third implicate smoking.

William Foege, who trained at the Harvard School, says of the tobacco lobby: "there are people knowing that their decision to advertise and promote tobacco is going to kill people... still willing to do it". But the tobacco industry is a major part of the economy, providing work for farmers, factory workers and shopkeepers. It pays dividends to its shareholders and provides enormous revenues for governments. In the United Kingdom, tobacco duty is worth about five pence on the standard rate of income tax.

What about those the economy does not benefit? To quote Foege again, "In almost any study you can find in any culture, mortality rates are higher for almost every category of disease in the poor as compared to the rich". There is something about poverty itself that makes it a health problem — a

sense of fatalism perhaps, and lack of a long-term optimistic perspective. People need an assured future. And it is difficult to get governments to do something about that. Either they will not because of their ideologies, or they cannot — which in many ways is worse.

After poverty, the next big problem facing the public-health professions is public perception. Letters to *The Times* in Victorian Britain fulminated against legal measures designed to prevent cholera. Fluoridation of water supplies to prevent dental decay seems to be an ideal public-health measure — effective, safe, cheap and easy to deliver. Yet only 11 per cent of the United Kingdom's water supplies are fluoridated. In the United States in the 1950s and 1960s, opponents of fluoridation were "rabid in their belief — and in their organised dissemination of the belief — that it was a Communist plot to deplete the brainpower and sap the strength of a generation of American children".

It is a risky world and risk sells newspapers. The media are old hands at whipping up a frenzy about health risks — asbestos, radiation, sex, drugs. But, as Henig says, the big risks are in people's lifestyles: smoking, driving, eating and drinking. These are personal choices, but made against a backdrop of national habits, and driven by business.

This book would not be complete without some success stories. Because it is about the Harvard School, most of the stories happened there. Some gestures are made to the rest of the world, for example John Snow's work on the cholera epidemic in London in the late 1840s. But no mention is made of Sir Edwin Chadwick, the lawyer and founding father of public health, who invented modern sewerage systems. The book's best story concerns Koch and his adversary Pettenkofer, who tried to demonstrate that Koch's microbe theory of cholera was rubbish by drinking a bottle of cholera vibrios — amazingly, with no harmful effects.

The book is good popular history. But we need more. We need the past to help us to decide on options for the future. Unfortunately the chapter 'Rethinking the Health Care System' ends disappointingly, saying that a consensus about how to go forward is unlikely. That may be true. But Henig may have fallen into the trap of thinking that health (and health-care) is one big issue rather than many smaller ones, some of which could be resolved.

The world is complicated and messy, and always will be. The idea of health is entangled with every other aspect of people's lives. But, if science is the art of the soluble and politics is the art of the possible, they should be able to achieve something between them. □

John Galloway is at the Eastman Dental Hospital, 256 Grays Inn Road, London WC1X 8LD, UK.

Three-dimensional structure of the tyrosine kinase c-Src

Wenqing Xu*, Stephen C. Harrison†‡§ & Michael J. Eck†

Laboratory of Molecular Medicine, Children's Hospital, and Departments of * Microbiology and Molecular Genetics, † Biological Chemistry and Molecular Pharmacology, and ‡ Pediatrics, Harvard Medical School, and § The Howard Hughes Medical Institute, 300 Longwood Avenue, Boston, Massachusetts 02115, USA

The structure of a large fragment of the c-Src tyrosine kinase, comprising the regulatory and kinase domains and the carboxy-terminal tail, has been determined at 1.7 Å resolution in a closed, inactive state. Interactions among domains, stabilized by binding of the phosphorylated tail to the SH2 domain, lock the molecule in a conformation that simultaneously disrupts the kinase active site and sequesters the binding surfaces of the SH2 and SH3 domains. The structure shows how appropriate cellular signals, or transforming mutations in v-Src, could break these interactions to produce an open, active kinase.

Src, the product of the first proto-oncogene to be characterized¹, is a protein switch. Its output is tyrosine phosphorylation, catalysed by a kinase domain; its inputs are multiple regulatory interactions, received by other parts of its polypeptide chain^{2,3}. The family of Src-related protein tyrosine kinases now includes nine members (Fyn, Yes, Fgr, Lyn, Hck, Lck, Blk and Yrk, in addition to Src itself). One or more of these proteins is present in every higher-animal cell type studied to date. Many, such as Lck, are specifically expressed in haematopoietic cells; Src itself shows a more widespread cellular distribution. Although not linked covalently to extracellular receptor domains, Src family members respond to a number of receptor-mediated signals, both by changes in kinase activity and by changes in intracellular localization².

Proteins of the Src family have a common domain organization, with each segment designated as a Src-homology (SH) region. The N-terminal segment includes the SH4 domain, which is a myristylation and membrane-localization signal, as well as a 'unique' domain of 50–70 residues that has no particular similarity to family members. The SH3, SH2 and catalytic (SH1) domains follow in order in the polypeptide chain; there is also a short, C-terminal 'tail', which includes a critical tyrosine residue. SH2 and

SH3 domains mediate protein–protein interactions in cellular signalling cascades, and are found in many proteins outside the Src family⁴. Extensive structural and functional studies of SH3 and SH2 domains have defined their specific molecular-recognition properties. SH3 domains are small, β -barrel modules that present a non-polar groove complementary to peptides in a polyproline-II conformation^{5,6}. Proline-rich sequences in target molecules mediate interactions with SH3-containing proteins⁷. SH2 domains bind polypeptide segments that contain a phosphotyrosine (pY)⁸, with significant context specificity, especially for the first three residues following the phosphotyrosine^{9–11}.

The SH3 and SH2 domains and the C-terminal tail all have roles in regulating Src kinase activity. The classically characterized v-src oncogene from avian retroviruses encodes a constitutively active kinase with a deleted tail¹². Mutation of the conserved tyrosine (Tyr 527 in avian c-Src) to phenylalanine also creates an active oncogenic protein¹³. It is now clear that phosphorylation of Tyr 527 by a specific kinase, Csk (ref. 14) inhibits Src catalytic activity by creating an intramolecular binding site for the Src SH2 domain^{15,16}. The interaction is believed to result in auto-inhibition by locking the molecule in an inactive state. Mutations in either the SH2 or the

Table 1 Data collection, phase determination, and refinement statistics

	Native	KAu(CN) ₂	PIP	KAu(CN) ₂ + PIP	Native 2
Resolution (Å)	25.0–2.0	25.0–2.5	25.0–2.45	25.0–2.0	20.0–1.5
R _{sym} (%)	5.3	4.2	3.9	4.9	4.9
Total observations	102,852	29,663	35,805	52,980	165,384
Unique reflections	25,752	14,690	15,631	24,653	48,728
Coverage (%)	90.0	89.7	90.1	80.4	66.1
R _{der.} (%)		24.9	15.3	28.0	
No. of sites		3	2	4	
Phasing power (25–2.0 Å, acent/cent)	1.79/1.35	0.63/0.55	1.50/1.19		
R _{cll} (%) (25–2.0 Å, acent/cent)		0.68/0.72	0.95/0.92	0.73/0.80	
Refinement (with native 2 data set)					
R _{free} (%) (20–1.5 Å)	26.4				
R _{cryst} (%) (20–1.5 Å)	21.0				
R.m.s.d.					
	Res. no.	Average B	Bonds (Å)	Angles (°)	B-values (Å ² bonded)
Protein	436	28.3	0.010	116	2.16
Water	490	39.4			

$R_{\text{sym}} = \sum_i |I_i - \langle I \rangle| / \sum_i I_i$, where I_i is the observed intensity, $\langle I \rangle$ is the average intensity of multiple observations of symmetry-related reflections. $R_{\text{der.}} = \sum_i |F_{\text{p}} - |F_{\text{p}}|| / \sum_i |F_{\text{p}}|$, where $|F_{\text{p}}|$ is the protein structure factor amplitude and $|F_{\text{p}}|$ is the heavy-atom derivative structure factor amplitude. Phasing power = r.m.s. ($|F_{\text{H}}|/E$), where $|F_{\text{H}}|$ is the heavy-atom structure factor amplitude and E is the residual lack of closure error. $R_{\text{cll}} = \sum_i |E_i| / \sum_i |F_{\text{H}}| - |F_{\text{p}}|$. Figure of merit = $(\sum P(\alpha) e^{i\alpha}) / \sum P(\alpha)$, in which α is the phase and $P(\alpha)$ is the phase probability distribution. $R = \sum_i |F_{\text{o}}| - |F_{\text{c}}| / \sum_i |F_{\text{o}}|$, where R_{free} is calculated for a randomly chosen 5% of reflections, R_{cryst} is calculated for the remaining 95% of reflections used for structure refinement.

SH3 domain can affect the stability of this 'closed', inactive conformation^{2,3,13}. Many, but not all, of these mutations are transforming in cell culture and tumorigenic in animals. In the 'open' state, phosphorylation of Tyr 416 in the kinase domain further enhances catalytic activity¹⁷.

To analyse the mechanism of this molecular switch, we have crystallized a large fragment of human c-Src containing the SH3,

SH2 and kinase domains and the C-terminal tail. We chose to prepare and use protein that is singly phosphorylated at Tyr 527, in order to select the closed, autoinhibited conformation. The structure, determined at 1.7 Å resolution, shows that there is indeed an intramolecular association of the SH2 domain with the phosphorylated Tyr 527. It also reveals how the SH3 domain contributes to the stability of the closed state, through an

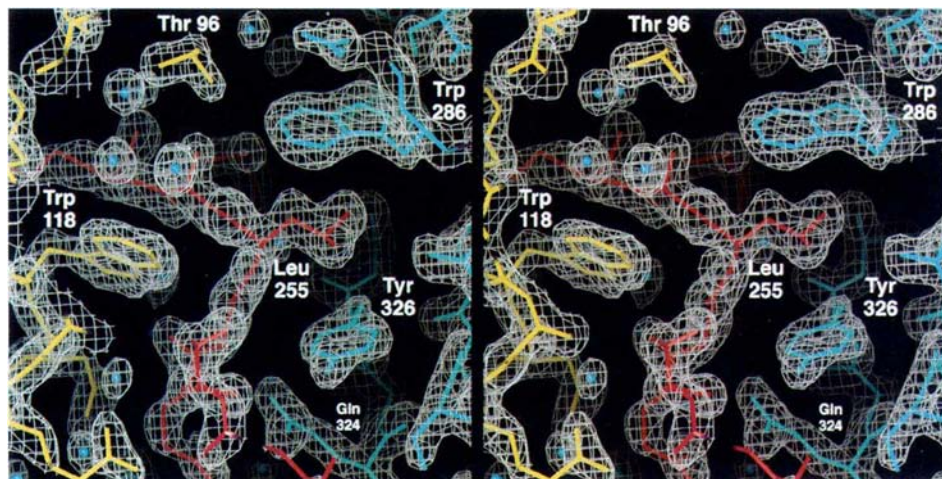


Figure 1 Stereo diagram showing the $2F_o - F_c$ electron density map and refined atomic model at the interface of the SH3, linker and kinase domains. The linker intercalates between the SH3 domain and the N-terminal lobe of the kinase domain, interacting extensively with both. The map is contoured at 1.5σ , and was

calculated using data between 20 and 1.5 Å resolution. The SH3 domain is yellow, the linker red, and the kinase domain light blue. Ordered water molecules are shown as blue spheres.

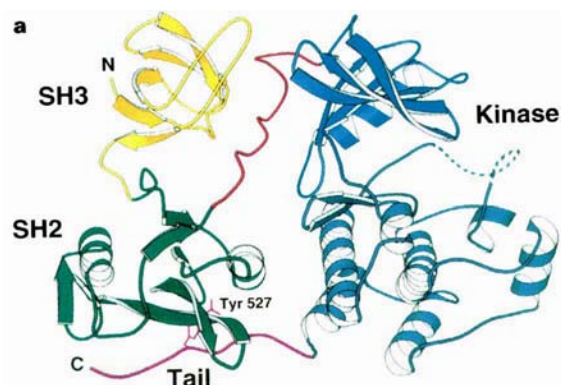
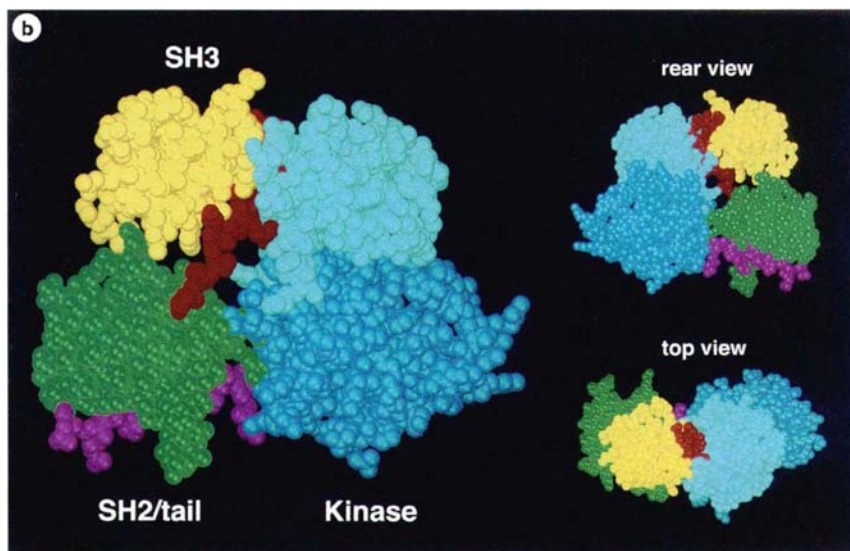


Figure 2 The 'closed' conformation of Tyr 527 phosphorylated c-Src. **a**, Ribbon diagram showing the structure and organization of domains. The SH3 and SH2 domains are coloured yellow and green, respectively. The linker connecting the SH2 and kinase domains (red) forms a polyproline II helix in complex with the SH3 domain. The N- and C-lobes of the catalytic domain are shown in cyan and blue. The disordered portion of the activation segment is shown as a dashed line. The phosphorylated tail (purple) is bound to the SH2 domain. **b**, Space-filling model. The SH2 domain makes only modest contact with the rest of the molecule. The RT and n-Src loops of the SH3 domain wrap around the linker to contact the N-lobe of the kinase in the front and rear, respectively.



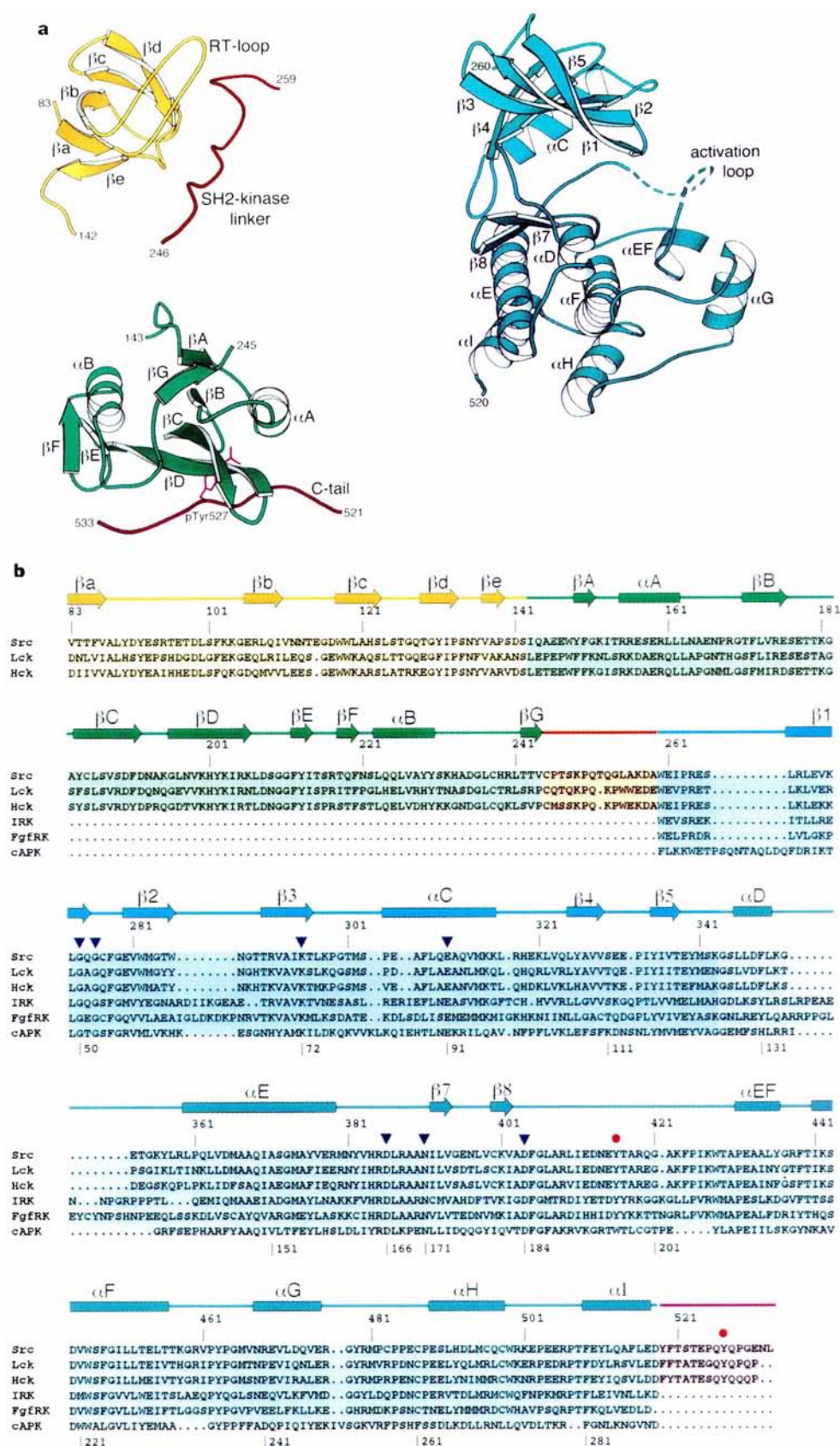


Figure 3 a, Ribbon diagrams of the SH3, SH2 and tyrosine kinase domains. Elements of secondary structure are labelled, using previously established nomenclature for the SH3, SH2 and catalytic domains. **b**, Structure-based sequence alignment, coloured by domain as in **a**. Sequences of the Src-family kinases Lck and Hck, the FGF (FgIRK) and insulin (IRK) receptor tyrosine kinases

and the cAMP-dependent serine kinase (cAPK) are shown. Active-site residues are indicated by inverted triangles; major regulatory phosphorylation sites by red dots (tyrosines 416 and 527 in c-Src). Residue numbers for c-Src and cAPK are shown above and below the sequence alignment, respectively.

unexpected interaction of the SH3 domain with the 'linker' that joins the SH2 and catalytic domains. The compactly organized, highly ordered domain assembly pushes the two lobes of the catalytic domain close together and enforces a conformation in the small lobe that disables the active site. Thus, in addition to dephosphorylation by tail-directed tyrosine phosphatases, competitive interactions with SH3 or SH2 ligands could destabilize the observed conformation. The tightly coupled interdomain contacts suggest that any one of these inputs is likely to produce the open, activated structure as its output.

Structure determination

We crystallized human c-Src (residues 86–536) in a closed, inactive conformation. The structure includes the SH3, SH2, tyrosine kinase, and 'C-terminal tail' domains of the protein; it is phosphorylated only on Tyr 527 in the regulatory tail. (For historical consistency, we refer to the structure using the numbering of chicken c-Src. Residues 86–536 of the human sequence correspond to residues 83–533 of chicken c-Src.) The divergent amino-terminal domain was deleted because it is not thought to be integrated into the closed form of the kinase¹⁸. The recombinant protein was produced using a baculovirus vector in Sf9 insect cells¹⁹. Isolation of a defined phosphorylation state, with stoichiometric phosphorylation of Tyr 527, was critical for crystallization and was accomplished with anion-exchange and phosphotyrosine-affinity chromatography (W. Xu *et al.*, manuscript in preparation). Mass analysis verified that the purified protein was monophosphorylated.

The structure was determined using conventional heavy atom/multiple isomorphous replacement methods, and refined to a crystallographic *R* value of 21% (*R*_{free} = 26.4%) using data between 20.0 and 1.5 Å resolution. Phase determination and crystallographic refinement statistics are shown in Table 1. A portion of the electron density map and the refined atomic model are shown in Fig. 1. The refined model includes all residues except 410–423 in the 'activation segment' of the catalytic domain, which are not visible in our density maps and therefore must be disordered.

Overall structure and domain organization

In its closed regulatory state, c-Src is a compact ensemble of four structural domains (Figs 2 and 3). The SH2 and SH3 domains lie beside the large and small lobes, respectively, of the tyrosine kinase domain, on the side opposite the catalytic cleft. The SH2 domain binds the phosphorylated C-terminal tail (pTyr 527), which extends from the base of the adjacent catalytic domain.

A 14-residue polypeptide linker, which joins the SH2 and catalytic domains, interacts along its course with both the SH3 domain and the small lobe of the kinase domain and serves as an adaptor to fit the two together. Although it contains only one proline, the linker (residues 246–259) adopts a polyproline type II (PPII) helical conformation in complex with the recognition surface of the SH3 domain (Fig. 4a). The SH3 domain and the small lobe of the kinase also contact each other directly.

The two lobes of the catalytic domain are opposed even more tightly than they are in the 'active' form of cyclic AMP-dependent protein kinase A (cAPK)^{20,21} (Figs 3a, 5a). Their close approach is accompanied by the displacement of an α -helix (helix C) in the small lobe from its position in active kinases, leading to rearrangement of residues that participate in catalysis (Fig. 5b). The conformation of the small lobe and its contacts across the catalytic cleft are strikingly similar to those seen in the uncomplexed (inactive) form of cyclin-dependent kinase 2 (Cdk2)^{22,23}. The conformation of the kinase domain appears to be stabilized by interactions of its small and large lobes with the SH3 and SH2 domains, respectively. There are direct contacts between the regulatory and catalytic domains, in addition to binding of the phosphorylated C-terminal tail to the SH2 domain and of the linker to the recognition surface of the SH3 domain. The structural organization of these interactions suggests a significant degree of cooperativity.

Domain architecture and interactions

Three-dimensional structures have been determined for the isolated SH3 and SH2 domains of Src and several closely related kinases. These domains do not show any significant conformational

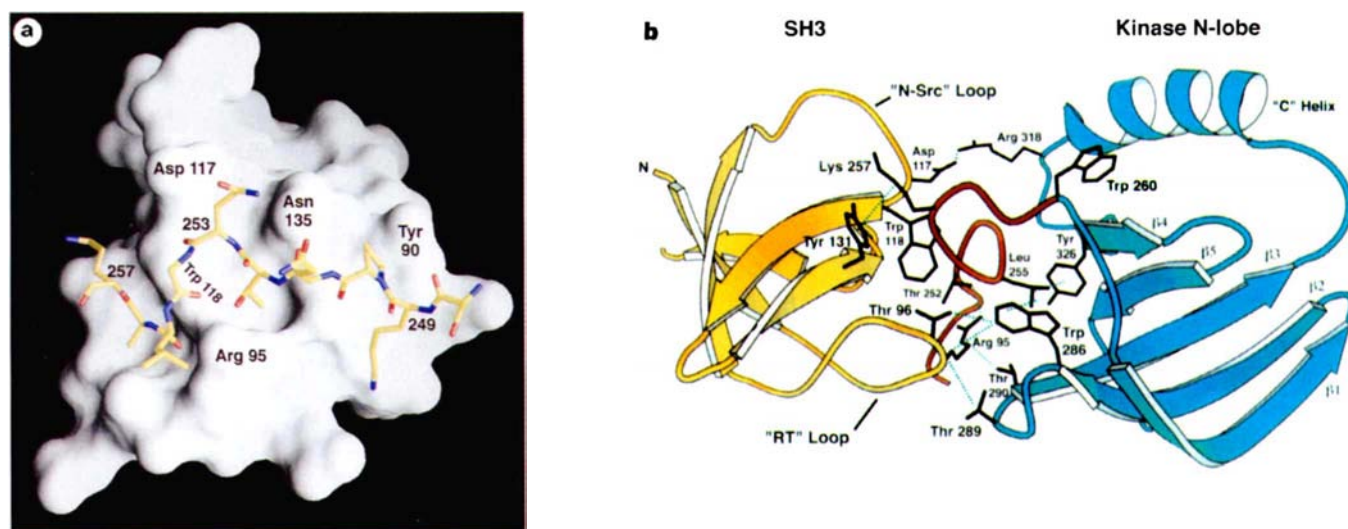


Figure 4 a, Molecular surface of the SH3 domain in complex with part of the linker segment. Residues in the domain are designated by name and number; those in the linker, by number only. The linker interacts in a class II orientation, with residues 249–253 in a polyproline type II (helical) conformation. C-terminal to Gln 253, the linker is not in a PPII conformation, but it continues to make extensive contact with the SH3 domain. Gly 254 introduces a kink, which allows the chain to arch over Trp 118. Leu 255 extends away from the surface of the SH3 domain and intercalates between Tyr 326 and Trp 286 in the kinase domain (see Fig. 1, and

panel **b**). The C-terminal residues of the linker (256–259) form a β -turn, which also packs between the SH3 and kinase domains. **b**, 'Top' view of the interactions between the SH3, linker and kinase domains. The linker (red) 'glues' the SH3 domain to the N-terminal lobe of the kinase. In addition, the RT and N-Src loops of the SH3 domain contact the kinase directly. Hydrogen-bond and van der Waals contacts are indicated by dashed green lines. Trp 260, the first residue in the kinase domain, packs against the 'C' helix and anchors the C-terminal portion of the linker.

rearrangements in our structure. The large and small lobes of the tyrosine kinase domain are also very similar to the known structures of the insulin²⁴ and fibroblast growth factor (FGF) receptor tyrosine kinases²⁵ and to several protein serine kinases²⁶.

The SH3 domain. The SH3 domain (residues 83–142) is a compact five-stranded β -sandwich (Fig. 3a). Its ligand-binding surface is formed by a cluster of hydrophobic residues and is flanked by the 'RT' and 'n-Src' loops, which connect strands β a and β b, and strands β b and β c, respectively. The RT loop is a site of activating mutations in v-Src^{27–29}; the n-Src loop contains a six-residue insertion in the more active neuronal isoform of Src³⁰. In our structure, the linker binds in the hydrophobic binding surface, and the RT and n-Src loops extend on either side of it to contact the catalytic domain (Figs 2b and 4). Residues 249–253 of the linker form a left-handed PPII helix and bind in a characteristic class II orientation, as seen with peptide ligands having PXXPR consensus^{31,32}. Proline 250, the only proline in this segment, packs between Tyr 90 and Tyr 136, the position usually assumed by the first proline in the class II binding sequence^{31,32}. Gln 253 occupies the other site that normally requires proline. Its long polar side chain cannot intercalate into the binding cleft, and so the course of the linker deviates from that of proline-rich peptides at this point (Fig. 4a).

Interaction of the n-Src loop with the catalytic domain is modest; Asp 117 forms a salt bridge with Arg 318 in the kinase domain. The RT loop makes a more extensive contact. In particular, Arg 95 and Thr 96 make van der Waals and hydrogen bond contacts with the β 2– β 3 loop in the catalytic domain, and their side chains are in hydrophobic contact with Trp 286, the last residue in strand β 2. The extended side chain of Arg 95 forms hydrogen bonds with Thr 252 and with the carbonyl of Leu 255 in the linker, and interacts through well-ordered water molecules with the hydroxyl groups of tyrosines 136 and 326 in the SH3 and catalytic domains, respectively (Fig. 4b).

Interaction between the SH3 and SH2 domains is minimal. Residues 142–146 at the SH2 amino terminus form a 3_{10} helical turn, which lies between small clusters of hydrophobic residues on each of the domains. The SH3/SH2 connection must be quite flexible; in a structure of an SH3/SH2 fragment of Lck, we observed a very different relative orientation of the two domains, even though the corresponding residues in Lck form an identical 3_{10} helical turn³³. The difference in observed orientation is due entirely to backbone rotation about the two residues just N-terminal to this turn (Asp 141 and Ser 142). We suggested that dimeric interactions in the SH3/SH2 crystal structure might serve as a model for the cooperative role of the SH3, SH2 and tail domains in regulation³³.

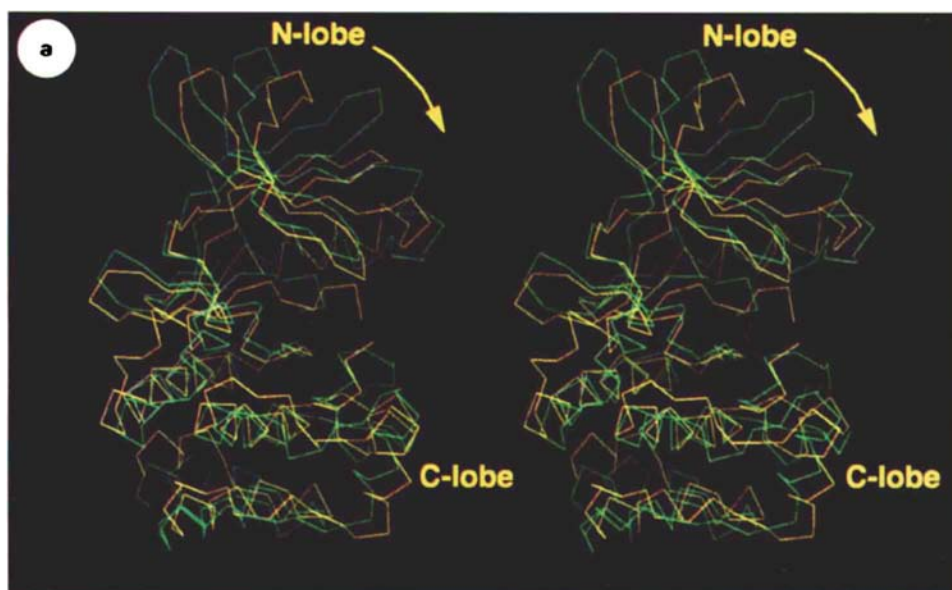
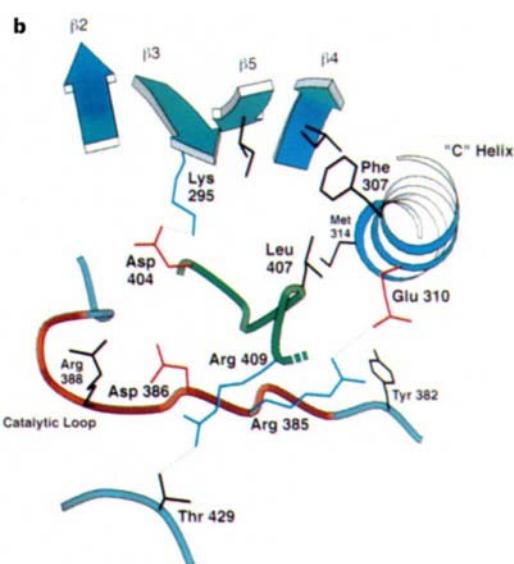


Figure 5 a, Alpha-carbon traces of the catalytic domains of c-Src (yellow) and cAPK (green), superimposed using the C-terminal lobes. In c-Src, interactions of the SH3 and SH2 domains rotate the N lobe by $\sim 9^\circ$ relative to its position in cAPK, closing the catalytic cleft. The break in the Src chain, seen at the right-hand edge of the central cleft, is the disordered part of the activation segment. **b**, The active site is partially disassembled. By analogy with cAPK, in an active conformation Glu 310 should contact Lys 295, positioning it to coordinate the α - and β -phosphates of ATP. In our structure, Glu 310 on helix C is rotated out of the catalytic conformation and is separated from Lys 295 by over 12 Å; its position in this inactive state is stabilized by an interaction with Arg 385. Leu 407, fixed in position by the anchoring of Arg 409, also enforces displacement of helix C. Phosphorylation of Tyr 416 may re-establish the catalytic configuration by coordinating Arg 385 and Arg 409, thereby relieving the electrostatic and steric barriers preventing helix C from assuming its active position. Other critical active-site residues (and in parentheses, their equivalents in cAPK) include the catalytic base, Asp 386 (Asp 166); Asp 404 (Asp 184), which coordinates a catalytic magnesium ion; and Arg 388 (Lys 168), which coordinates the γ -phosphate.



Our structure shows that this model is not relevant to the inactive state.

The SH2 domain. The conserved SH2 domain fold includes a central four-stranded β -sheet and two α -helices, which pack on either side of it (Fig. 3a). Numerous X-ray and NMR structures of SH2 domains in complex with phosphotyrosine-containing peptides show similar modes of recognition—phosphopeptides bind in an extended conformation across the surface of the domain, roughly perpendicular to the β -D edge of the central sheet³⁴. Phosphotyrosine is bound in a pocket on one side of the central sheet, and in high-affinity complexes, residues C-terminal to phosphotyrosine bind in a pocket or groove on the opposite side. In Src-family SH2 domains, a hydrophobic pocket recognizes a leucine or isoleucine residue at position pY + 3 in high-affinity peptides^{10,11}. The phosphorylated tail in our structure binds with the SH2 domain in a manner suggestive of a low-affinity interaction³⁴, consistent with experimental measurements of the affinity of the isolated Src SH2 domain for phosphopeptides modelled on the C-terminal tail³⁵. The conformation and interactions of residues pY - 1 to pY + 2 (Gln-pTyr-Gln-Pro) in the tail are the same as those made by the corresponding residues in high-affinity polyoma middle-T sequence in previous Src and Lck SH2 structures^{10,11}, but outside this region the tail is poorly ordered and does not appear to make specific interactions. In particular, there is no side chain occupying the pY + 3 pocket.

The phosphorylated tail is a short but flexible tether that anchors the SH2 domain. Apart from the tail, the SH2 domain makes contact with the catalytic domain only along its A-helix, which runs roughly antiparallel to the E-helix of the kinase domain. The helices are not closely packed. Their juxtaposed surfaces are electrostatically complementary, however, and several charged and polar side chains interact across their interface, which also contains a number of bound water molecules.

Tyrosine kinase. The c-Src catalytic domain has an unembellished protein kinase fold, comprising a smaller, N-terminal lobe connected by a flexible 'hinge' to a larger, C-terminal lobe²⁰. The N-terminal lobe is a five stranded antiparallel β -sheet, with a single helix (helix C) connecting strands β 3 and β 4; the C-terminal lobe is mostly α -helical (Fig. 3a). With the exception of helix C, the small and large lobes of the Src kinase domain superimpose well on the corresponding parts of cAPK; for example, 134 α -carbon atoms in the core of the large domain superimpose with an r.m.s.d. of 1.7 Å.

The two lobes of the catalytic domain approach each other closely. Superposition on the active conformation of cAPK, using only residues of the large lobe, shows that the small lobe in Src is rotated towards the large lobe by 9° relative to cAPK, creating an even more occluded interface (Fig. 5a). Superposition of the small lobes show that helix C in Src is displaced from this interface by ~5 Å and rotated away from it relative to its position in cAPK (Figs 3b, 5b). A precisely similar rearrangement is seen in the inactive form of Cdk2 (ref. 22). Helix C contains the conserved residue Glu 310 (Fig. 5b). Structures of kinases in their active conformations show that the side chain of this glutamate projects into the catalytic cleft to form a salt bridge with the equivalent of Lys 295, an important ligand for the α - and β -phosphates of ATP. In our structure (and in Cdk2), the glutamate faces outwards. It forms an alternative salt bridge with Arg 385, while Lys 295 interacts instead with Asp 404. The displaced helix C bears against Leu 407 near the beginning of the so-called 'activation segment'; this interaction prevents the helix from assuming its catalytic position (Fig. 5b). Leu 407 is the second residue in a β -turn, which is firmly anchored by insertion of the side chain of Arg 409 into a pocket near Thr 429. Residues following Arg 409 are disordered. In Cdk2, it is also a structured region near the beginning of the activation segment that displaces helix C (the PSTAIRE helix)²². Thus the same local inactivation mechanism is used in both Src and Cdk2. However, the molecular interactions controlling this local switch are

quite different. Cdk2 is activated by interaction with cyclin A, which binds helix C and pushes it into its active conformation²³. In Src, restoration of helix C to its active position is probably accomplished by phosphorylation of Tyr 416, and perhaps by disruption of the intramolecular SH3 and SH2 interactions, as discussed below.

Activation of the Src kinase

Reorientation of helix C to create a functional kinase would require a conformational change in the activation segment to remove interference from Leu 407. In a recent structure of the kinase domain of Lck, activated by phosphorylation at the position equivalent to Tyr 416, the phosphotyrosine forms salt bridges with the equivalents of both Arg 385 and Arg 409 (ref. 36). The homologous interactions in activated Src would necessarily expel Arg 409 from the pocket that it occupies in our structure, leading to rearrangement and ordering of the entire activation segment (Fig. 5b). Moreover, the side chain of Arg 385 would move away from Glu 310. Helix C would then be free to shift back to the position required for an active catalytic site.

How is this local switch, mediated by a concerted set of salt-bridge rearrangements, coupled to the global regulatory switch, mediated by the SH3 and SH2 domains? The structure described here suggests that the observed interactions of the regulatory domains with the linker, the C-terminal tail, and the kinase domain itself, all serve to push the lobes of the kinase together, reducing access to the catalytic cleft and forcing helix C to shift outwards. The Hck structure, reported elsewhere in this issue³⁷, shows that very similar SH3 and SH2 contacts are also compatible with some opening up of the lobes of the catalytic domain, perhaps due to the presence of quercetin in those crystals, which binds at the adenine site. Nonetheless, helix C is still in its displaced (inactive) position in the tail-phosphorylated Hck. Src phosphorylated on Tyr 527 in addition to Tyr 416 retains about 20% of its kinase activity³⁸. Catalysis must require the 'active' position of helix C. It is not yet clear, however, whether catalysis can occur with the regulatory domains clamped in place, or if it requires release of the C-terminal tail from the SH2 domain and the linker from the SH3. The structures of other phosphorylated states of Src will be useful for answering these questions.

The regulatory switch

The structure and regulation of Src and its family members have been analysed extensively by site-specific mutagenesis and by comparison of c-Src and v-Src². Introduction of isolated mutations into the SH3, SH2, kinase and tail regions of c-Src can be sufficient to create a constitutively active, and potentially transforming protein. In v-Src from Rous sarcoma virus, an unrelated 12-residue fragment replaces the last 19 residues of c-Src¹², removing the final turn of the 'I' helix and the entire regulatory tail. Various strains of the virus bear additional point mutations in the v-Src SH3 and catalytic domains³³. The tail mutation alone activates the kinase and leads to a transforming protein, suggesting that the conformation we observe is destabilized in the absence of the SH2/tail interaction. Some of the point mutations are in solvent-exposed residues, and these would not be expected to affect the structure or activity of the kinase, but others, including tryptophan swapped for arginine at position 95 (Arg95Trp), Thr97Ile, Asp117Asn and Thr338Ile, are found at domain interfaces and could further destabilize the closed conformation. Introduced on their own, substitutions for Arg 95 and Thr 96 in the RT loop are activating and partially transforming^{27,28}, and we would indeed expect them to disrupt the interaction of the SH3 domain with the linker and with the N-terminal lobe of the catalytic domain, thereby destabilizing the closed conformation. Likewise, we would expect mutation of Asp 117 to Asn to disrupt the interaction of the n-Src loop with the catalytic domain. Transformation by the isolated Thr 338 → Ile mutation is intriguing²⁷. This residue is at the back of the

nucleotide-binding pocket, near the interdomain hinge. Introduction of isoleucine at this position would disrupt local hydrogen-bond interactions and might also require a small domain rotation, which could destabilize the closed conformation.

Spontaneous single-site mutations of Glu 378 to glycine, and of Ile 441 to phenylalanine, confer transforming activity on c-Src³⁹. These residues pack together in the three-dimensional structure, just under Tyr 382, which helps to stabilize Glu 310 in its inactive position (Fig. 5b). Thus, disruption of this interaction may trigger activation of the kinase.

The interdependence of Tyr 527 phosphorylation and the SH3 and SH2 domains in regulation has been explicitly examined in a yeast system by coexpressing various Src mutants and Csk. These studies demonstrate that down-regulation of the kinase by phosphorylation of the tail requires both the SH3 and SH2 domains and, in particular, an intact ligand-binding surface and RT loop in the SH3 domain^{29,40–42}. The present structure shows that the SH3 domain binds the linker and the N-lobe of the kinase with this surface, and in turn positions the SH2 domain to interact with the kinase and phosphorylated tail.

A number of cellular inputs, including tail dephosphorylation or apposition of a high-affinity ligand for the SH2 or the SH3 domain⁵², could produce a transition to the open, active conformation. The flexibility of the SH3/SH2 interaction and the long linker between the SH2 and kinase domains suggest that the open conformation might be relatively floppy, with little interaction among the component domains. The regulatory domains would be free in such a state to bind their cognate cellular targets and to direct activated c-Src to its appropriate substrate and subcellular location. □

Methods

Protein purification. Insect cells bearing c-Src(Δ M85)¹⁹ were lysed in 150 mM NaCl, 25 mM HEPES, pH 7.6, and 5 mM DTT. The lysate was cleared by ultracentrifugation, and c-Src(Δ N85) was isolated by sequential column chromatography on DEAE-Sephacrose CL-6b, γ -aminophenyl ATP-Sephacrose⁴³, and Superdex-200 (Pharmacia). Phosphorylation states were defined by mass spectroscopy, routinely identified by native PAGE electrophoresis (Phastegel, Pharmacia), and separated by phosphotyrosine and Q-Sepharose HP chromatography. Purified protein was maintained in a storage buffer containing 20 mM HEPES, pH 7.6, 0.1 M NaCl and 5 mM DTT. The yield of Tyr 527-phosphorylated material was increased by incubating non-phosphorylated protein (1 mg ml⁻¹) in storage buffer with purified recombinant Csk (1 mg ml⁻¹), in the presence of 10 mM MnCl₂ and 1.0 mM ATP.

Crystallization and data collection. Src crystals were grown in hanging drops by combining 1 μ l protein solution (15 mg ml⁻¹ protein in storage buffer) with 1 μ l reservoir solution (50 mM PIPES, pH 6.5, 0.8 M sodium tartrate, 20 mM DTT). Crystals grow to maximum size (0.25 $\times$ 0.25 $\times$ 0.6 mm³) in about a week at room temperature. Different crystal forms can grow in the same drop; the form we used belongs to space group P2₁2₁2₁, and has cell dimensions $a = 51.76$ Å, $b = 87.38$ Å, $c = 101.30$ Å, with one molecule per asymmetric unit. For derivatization and cryogenic data collection, crystals were transferred stepwise into 50 mM PIPES, pH 6.5, 1.15 M sodium tartrate, 15% glycerol and 0.1 M NaCl, and allowed to stabilize for at least 24 h before transfer to a final stabilizing solution containing 20% glycerol, 20 mM PIPES, pH 6.5, 1.15 M sodium tartrate and 0.1 M NaCl.

Diffraction data from native and derivative crystals were recorded with a Mar Research image plate scanner mounted on an Elliot GX-13 rotating anode source with mirror optics, and integrated and scaled with the programs DENZO and SCALEPACK⁴⁴ or XDS and XSCALE⁴⁵ (Table 1). A high-resolution native data set, complete to ~ 2.1 Å but $\sim 55\%$ complete between 2.1 and 1.5 Å, was recorded (from the same native crystal used for phasing) on an ADSC IK CCD detector at the CHESS F-1 beamline (Cornell University). The synchrotron data set was merged with the native 1 data set to create native 2, the data set used in refinement. We estimate an 'effective' resolution of 1.69 Å, based on the number of reflections observed with $I/\sigma > 2$.

Structure determination and refinement. The structure was determined by conventional multiple isomorphous replacement (MIR) using KAu(CN)₂

(1 mM, 2 days), di- μ -iodobis (ethylene diamine) diplatinum (II) nitrate (PIP) (50 μ M, 18 h) and a double-soak (KAu(CN)₂ and PIP) as derivatives (Table 1). Heavy-atom positions were located by Patterson and difference Fourier methods with the CCP4 program package⁴⁶. Heavy-atom parameters were refined and phases were calculated with MLPHARE (overall figure of merit was 0.483 at 25.0–2.0 Å resolution). The initial MIR map was improved with real-space density modification using the program DM⁴⁶. Skeletonization of the improved map with BONES⁴⁷ allowed us to fit models of the Lck SH3 domain¹¹, v-Src SH2 domain¹⁰, and N and C domains of the insulin-receptor kinase²⁴ as rigid bodies. This model was rebuilt to reflect the sequence of human c-Src and refitted to the map using program O⁴⁷. After an initial cycle of crystallographic refinement using XPLOR⁴⁸, the linker and tail regions were built. The model was refined with simulated annealing and positional refinement and manual rebuilding using the programs XPLOR⁴⁸ and O⁴⁷. Simulated-annealing omit maps were computed to check the conformation of C-terminal tail. Water molecules were added with the aid of the program ARP⁴⁹. Refinement statistics are given in Table 1. There are no residues with disallowed main-chain torsion angles. The refined model includes residues 83–409, 424–533, the N-terminal methionine, and 490 water molecules. Phe 424 is modelled as an alanine. The main-chain density is well-defined except for residues 521–525 and 530–533 in the 'tail' region which are poorly ordered and probably adopt multiple conformations.

Illustrations. Figures 1 and 2b were prepared with O⁴⁷; Figs 2a, 3a, 4b and 5b with MOLSCRIPT⁵⁰. The program GRASP⁵¹ was used to create Figs 4a and 5a.

Received 13 December 1996; accepted 22 January 1997

- Bishop, J. Viral Oncogenes. *Cell* **42**, 23–28 (1985).
- Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of src. *Biochim. Biophys. Acta* **1287**, 121–149 (1996).
- Superti-Furga, G. & Courtneidge, S. A. Structure–function relationships in Src family and related protein tyrosine kinases. *Bioessays* **17**, 321–330 (1995).
- Pawson, T. Protein modules and signalling networks. *Nature* **373**, 573–580 (1995).
- Yu, H. *et al.* Solution structure of the SH3 domain and Src and identification of its ligand-binding site. *Science* **258**, 1665–1668 (1992).
- Musacchio, A., Saraste, M. & Willmanns, M. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nature Struct. Biol.* **1**, 546–551 (1994).
- Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. Identification of a ten amino acid proline-rich SH3 binding site. *Science* **259**, 1157–1161 (1993).
- Mayer, B. J., Jackson, P. K. & Baltimore, D. The noncatalytic src homology region 2 segment of *abl* tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. *Proc. Natl. Acad. Sci. USA* **88**, 627–631 (1991).
- Songyang, Z. *et al.* SH2 domains recognize specific phosphopeptide sequences. *Cell* **72**, 767–778 (1993).
- Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D. & Kuriyan, J. Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell* **72**, 779–790 (1993).
- Eck, M. J., Shoelson, S. E. & Harrison, S. C. Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56^{lck}. *Nature* **362**, 87–91 (1993).
- Takeya, T. & Hanafusa, H. Structure and sequence of the cellular gene homologous to the RSV src gene and the mechanism for generating the transforming virus. *Cell* **32**, 881–890 (1983).
- Hunter, T. A tail of two src's: mutatis mutandis. *Cell* **49**, 1–4 (1987).
- Nada, S., Okada, M., MacAuley, A., Cooper, J. A. & Nakagawa, H. Cloning of a complementary DNA for a protein-tyrosine kinase that specifically phosphorylates a negative regulatory site of p60c-src. *Nature* **351**, 69–72 (1991).
- Matsuda, M., Mayer, B. J., Fukui, Y. & Hanafusa, H. Binding of transforming protein, P47gag-crk, to a broad range of phosphotyrosine-containing proteins. *Science* **248**, 1537–1539 (1990).
- Roussel, R. R., Brodeur, S. R., Shalloway, D. & Laudano, A. P. Selective binding of activated pp60c-src by an immobilized synthetic phosphopeptide modeled on the carboxyl terminus of pp60c-src. *Proc. Natl. Acad. Sci. USA* **88**, 10696–10700 (1991).
- Cooper, J. A. & Howell, B. The when and how of Src regulation. *Cell* **73**, 1051–1054 (1993).
- Koegl, M., Courtneidge, S. A. & Superti-Furga, G. Structural requirements for the efficient regulation of the Src protein tyrosine kinase by Csk. *Oncogene* **11**, 2317–2329 (1995).
- Ellis, B. *et al.* Purification and characterization of deletion mutations of pp60c-src tyrosine kinase. *J. Cell. Biochem. (suppl.)* **18B**, 276 (1994).
- Knighton, D. R. *et al.* Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 407–414 (1991).
- Madhusudan *et al.* cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphotransfer. *Protein Sci.* **3**, 176–187 (1994).
- De Bondt, H. L. *et al.* Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595–602 (1993).
- Jeffrey, P. D. *et al.* Mechanism of CDK activation revealed by the structure of a cyclinA–CDK2 complex. *Nature* **376**, 313–320 (1995).
- Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746–754 (1994).
- Mohammadi, M., Schlessinger, J. & Hubbard, S. R. Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. *Cell* **86**, 577–587 (1996).
- Johnson, L. N., Noble, M. E. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149–158 (1996).
- Kato, J. Y. *et al.* Amino acid substitutions sufficient to convert the nontransforming p60c-src protein to a transforming protein. *Mol. Cell. Biol.* **6**, 4155–4160 (1986).
- Potts, W. M., Reynolds, A. B., Lansing, T. J. & Parsons, J. T. Activation of pp60c-src transforming potential by mutations altering the structure of an amino terminal domain containing residues 90–95. *Oncogene Res.* **3**, 343–355 (1988).
- Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S. A. & Draetta, G. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. *EMBO J.* **12**, 2625–2634 (1993).

30. Levy, J. B. & Brugge, J. S. Biological and biochemical properties of the c-src+ gene product overexpressed in chicken embryo fibroblasts. *Mol. Cell. Biol.* **9**, 3332–3341 (1989).
31. Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science* **266**, 1241–1247 (1994).
32. Lim, W. A., Richards, F. M. & Fox, R. O. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372**, 375–379 (1994).
33. Eck, M. J., Atwell, S. K., Shoelson, S. E. & Harrison, S. C. Structure of the regulatory domains of the Src-family tyrosine kinase Lck. *Nature* **268**, 764–769 (1994).
34. Kuriyan, J. & Cowburn, D. Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct.* (in the press).
35. Payne, G., Shoelson, S. E., Gish, G. D., Pawson, T. & Walsh, C. T. Kinetics of p56lck and p60src Src homology 2 domain binding to tyrosine-phosphorylated peptides determined by a competition assay or surface plasmon resonance. *Proc. Natl Acad. Sci. USA* **90**, 4902–4906 (1993).
36. Yamaguchi, H. & Hendrickson, W. A. Structural basis for activation of the human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* **384**, 484–489 (1996).
37. Sicheri, F., Moarefi, I. & Kuriyan, J. *Nature* **385**, 602–609 (1997).
38. Boerner, R. J. et al. Correlation of the phosphorylation states of pp60 c-src with tyrosine kinase activity: the intramolecular pY530-SH2 complex retains significant activity if Y419 is phosphorylated. *Biochemistry* **35**, 9519–9525 (1996).
39. Levy, J. B., Iba, H. & Hanafusa, H. Activation of the transforming potential of p60c-src by a single amino acid change. *Proc. Natl Acad. Sci. USA* **83**, 4228–4232 (1986).
40. Murphy, S. M., Bergman, M. & Morgan, D. O. Suppression of c-Src activity by C-terminal Src kinase involves the c-Src SH2 and SH3 domains: analysis with *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13**, 5290–5300 (1993).
41. Okada, M., Howell, B. W., Broome, M. A. & Cooper, J. A. Deletion of the SH3 domain of Src interferes with regulation by the phosphorylated carboxyl-terminal tyrosine. *J. Biol. Chem.* **268**, 18070–18075 (1993).
42. Erpel, T., Superti-Furga, G. & Courtneidge, S. A. Mutational analysis of the Src SH3 domain: the same residues of the ligand binding surface are important for intra- and intermolecular interactions. *EMBO J.* **14**, 963–975 (1995).
43. Haystead, C. M., Gregory, P., Sturgill, T. W. & Haystead, T. A. Gamma-phosphate-linked ATP-sepharose for the affinity purification of protein kinases. Rapid purification to homogeneity of skeletal muscle mitogen-activated protein kinase kinase. *Eur. J. Biochem.* **214**, 459–467 (1993).
44. Otwinowski, Z. in *Proceedings of the CCP4 Study Weekend* (eds Sawyer, L., Isaacs, N. & Burley, S.) 56–62 (SERC Daresbury Laboratory, Daresbury, UK, 1993).
45. Kabsch, W. Evaluation of single crystal diffraction data from a position sensitive detector. *J. Appl. Crystallogr.* **21**, 916–924 (1988).
46. Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–776 (1994).
47. Jones, T. A., Bergdoll, M. & Kjeldgaard, M. in *Crystallographic Computing and Modeling Methods in Molecular Design* (eds Bugg, C. & Ealick, S.) (Springer, New York, 1989).
48. Brunger, A. T. *X-PLOR Version 3.0: A System for Crystallography and NMR* (Yale University Press, New Haven, CT, 1992).
49. Lamzin, V. S. & Wilson, K. S. Automated refinement of protein models. *Acta Crystallogr. D* **49**, 129–147 (1993).
50. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
51. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).
52. Alexandropoulos, K. & Baltimore, D. Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel p130Cas-related protein. *Sin. Genes Dev.* **10**, 1341–1355 (1996).

Acknowledgements. We thank B. Ellis, L. Overton, C. Hoffman, R. Boerner, K. Blackburn, W. B. Knight, M. Milburn and M. Luther of Glaxo-Wellcome for the c-Src(Δ N85) insect cell pellet, for mass spectroscopic analysis, and for discussions; M. Lawrence and the staff at CHESS for help with synchrotron data collection; P. Cole for purified recombinant Csk; K. Svenson for technical assistance; R. Nolte for help with crystallographic computation; J. Kuriyan and co-workers, for communicating results prior to publication; and T. Sweeney for help in preparing the manuscript. W.X. is supported by the Irvington Institute for Medical Research. This work was funded in part by a grant from BASF Bioresearch Corporation to M.J.E. S.C.H. is an investigator in the Howard Hughes Medical Institute. M.J.E. is the recipient of a Burroughs-Wellcome Fund Career award in the Biomedical Sciences.

Correspondence and requests for materials should be addressed to M.J.E. Coordinates have been deposited at the Brookhaven Protein Data Bank, accession code 1FMK, and are also available by e-mail from eck@rascal.med.harvard.edu.

Crystal structure of the Src family tyrosine kinase Hck

Frank Sicheri[†], Ismail Moarefi^{*†} & John Kuriyan^{*}

Laboratories of Molecular Biophysics, and ^{*} Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

[†] These authors contributed equally to this work.

The crystal structure of the haematopoietic cell kinase Hck has been determined at 2.6/2.9 Å resolution. Inhibition of enzymatic activity is a consequence of intramolecular interactions of the enzyme's Src-homology domains SH2 and SH3, with concomitant displacement of elements of the catalytic domain. The conformation of the active site has similarities with that of inactive cyclin-dependent protein kinases.

The Src-family kinases, named after the *src* oncogene of Rous sarcoma virus, are a closely related group of non-receptor tyrosine kinases that play critical roles in eukaryotic signal transduction. The viral and cellular forms of the *src* gene, *v-src* and *c-src*, were the first oncogene–proto-oncogene pair to be identified (reviewed in ref. 1). A critical difference between *v-Src* and *c-Src* is the loss in the former of a regulatory tyrosine residue that is located in the carboxy-terminal tail of *c-Src* (Tyr 527). Phosphorylation of Tyr 527 reduces the tyrosine kinase activity of *c-Src*, and without this inhibitory control the viral form of the protein greatly elevates cytoplasmic levels of tyrosine phosphorylation and is a potent cellular transforming agent¹. With the aim of understanding the molecular basis for this critical signalling mechanism we have pursued a crystallographic investigation of the Src-family member Hck (haematopoietic cell kinase) in the auto-inhibited form.

The Src-family members share a common regulatory mechanism, but differ in cellular expression and localization. Nine Src-family tyrosine kinases have been identified (Src, Lck, Hck, Fyn, Fgr, Yes, Blk, Lyn and Yrk)¹. *c-Src* is widely expressed and phosphorylates a wide range of substrates, whereas Lck plays a more restricted but critical role in T-cell signalling. Hck, the subject of this study, is expressed in lymphoid and myeloid cells^{2,3}, and is bound to B-cell receptors in unstimulated B cells. Knockout studies in mice have shown that simultaneous deletion of the genes for *hck* and *src*, or *hck* and *fgr*, leads to severe developmental anomalies and impaired immunity^{4,5}.

The highly conserved regulatory apparatus of the Src family members consists of two peptide-binding modules, the Src-homology domains SH2 and SH3 (refs 6, 7). These modules bind to targets containing phosphotyrosines and polyproline type II helices, respectively, and mediate the formation of protein–protein complexes during signalling⁸. Interactions between the SH2 domain and a C-terminal phosphotyrosine residue in the Src kinases (Tyr 527 in *c-Src*) results in repression of catalytic activity, with additional inhibitory interactions provided by the SH3 domain. The inhibitory phosphorylation at Tyr 527 is mediated by a distinct tyrosine kinase, Csk (*c-Src* kinase), whereas autophosphorylation at another tyrosine residue (Tyr 416 in *c-Src*), located within the 'activation segment' of the catalytic domain, is required for catalytic activity¹.

The catalytic domain alone is functional as a tyrosine kinase, but the SH2 and SH3 domains are required for full biological activity. Certain mutations in these domains lead to host cell-dependent phenotypes, indicating that they play a dual role in the Src kinases¹. The SH2 and SH3 domains are required for inhibition of the enzyme, but once released from that role they function to target the kinase to specific substrates.

We now describe the crystal structure, determined at 2.6/2.9 Å resolution, of the downregulated form of the haematopoietic cell kinase, Hck. The SH2 domain, long implicated in regulating enzyme activity, is bound to the C-terminal phosphorylated tail, but is distant from the active site. Unexpectedly, the linker con-

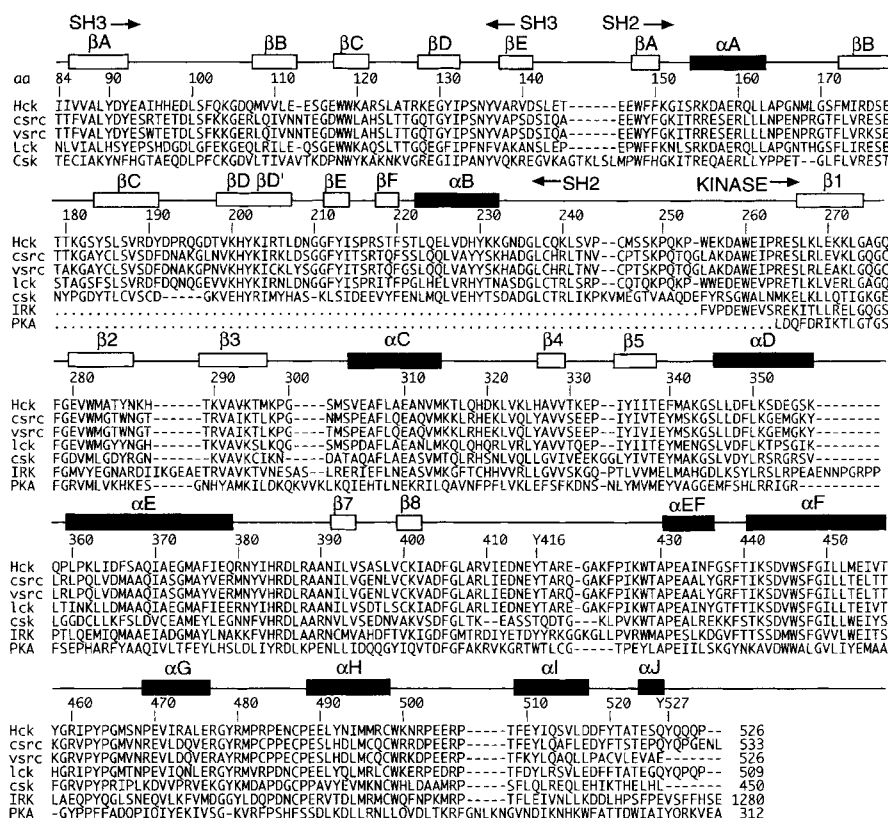


Figure 1 Sequence alignment of human Hck (residues 81–526), chicken c-Src (residues 84–533), chicken v-Src (residues 84–526), human Lck (residues 64–509), human Csk (residues 12–450), human IRK (residues 984–1,280) and mouse PKA (residues 40–312). Secondary structural elements and domain boundaries are indicated above the aligned sequences, with β -strands and α -helices indicated by white and black boxes, respectively. In **a**, the catalytic segment and strands β 7– β 8 are orange.

Table 1 Data collection, structure determination and refinement statistics

$I/\sigma > 0$	Resolution (Å)	Reflections total/unique	Completeness (%)	R_{sym}^* (%)	$R_{\text{iso}}^\dagger$ (%)	Sites (No.)	Phasing power ‡ (3.5 Å/3.2 Å)	FOM § (3.5 Å/3.2 Å)
Native data								
Native (Quercetin) $\parallel$	3.0	38,5451/24,477	98.1(97.2)	7.9(14.5)				
Native (AMP-PNP) $\parallel$	2.6	34,6924/34,027	96.8(92.8)	5.3(18.9)				
MIRAS analysis								
EMP	3.2	158,183/19,974	98.0(96.3)	8.3(22.4)	29.3	4	1.61/1.54	0.45/0.39
Thimerosal	3.2	215,839/20,391	99.8(98.7)	8.3(22.7)	31.4	8	1.69/1.60	
Mersalyl Acid	3.3	110,963/15,993	86.1(90.7)	11.7(34.4)	37.9	4	0.86/0.86	
MAD analysis $^\#$								
$\lambda_1 = 1.0063$ Å	3.1	196,158/21,812	96.9(96.3)	7.3(27.9)		8	1.77/1.55	0.42/0.37
$\lambda_2 = 1.0093$ Å	3.3	177,533/18,501	97.9(98.2)	8.2(31.4)	12.8	8	1.18/1.17	
MIRAS-MAD combined								0.57/0.51
		Reflections	Completeness		R_{factor}	R_{free}^{**}		
Refinement (quercetin complex)								
3.0–18 Å	$I/\sigma > 3$	23,857	95.6(86.8)		22.9	31.5		
3.0–18 Å	all	24,477	98.1(97.2)		23.3	32.3		
Refinement (AMP-PNP complex)								
2.9–18 Å $\parallel$	$I/\sigma > 3$	24,109	88.5(67.1)		22.7	30.6		
2.6–18 Å $\parallel$	$I/\sigma > 3$	29,868	79.4(47.1)		23.7	31.2		
r.m.s. deviations of model (AMP-PNP/quercetin)								
Bond length (Å)		0.016/0.018		Total non-hydrogen atoms		7,038/7,020		
Bond angle (°)		3.2/3.7		Calcium ions		4/2		
Ave. B -factor (Å ²)		42.7/57.5		Water molecules		0/2		
r.m.s.d. in B -factor for bonded atoms (Å ²)		2.0/1.9						

$R_{\text{sym}} = 100 \times \sum_i \sum_j |I_{h,j} - \langle I_h \rangle| / \sum_i \sum_j I_{h,j}$. For R_{sym} and completeness, the numbers in parentheses refer to data in the shell of highest resolution.
 $R_{\text{iso}} = \sum_i |F_{\text{native}} - F_{\text{derivative}}| / \sum_i F_{\text{native}}$. For the MAD analysis, R_{iso} is calculated between data sets at two wavelengths. Wavelength 1 is used as reference data set, F_{native} in this calculation, and $F_{\text{derivative}}$ refers to data at wavelength 2.
 ‡ Phasing power is r.m.s. ($|F_h|/E$) where the subscript 'h' represents 'heavy-atom' and E is the residual lack of closure.
 § FOM (mean figure of merit) = $\langle |P(\alpha)|^2 / P(\alpha) \rangle$, where α is the phase and $P(\alpha)$ is the phase probability distribution.
 $^\parallel$ Statistics reflect the merging of diffraction data from two crystals.
 $^\parallel$ The intensities are anisotropic, as indicated by the completion statistics obtained when using an $I/\sigma > 3$ cut off.
 $^\#$ Thimerosal derivative.
 ** Free R -value was calculated with 5% of the data.

necting the SH2 domain to the catalytic domain forms a polyproline type II helix, which serves as a docking site for the SH3 domain. The role of the SH2 domain appears to be to position the SH3 binding site, and it is the SH3 interaction that provides a direct connection to the regulation of kinase activity. The mechanism of inactivation of the catalytic domain has elements in common with that of the cyclin-dependent protein kinases.

Structure determination

Human Hck (residues 78 to 526, excluding the amino-terminal membrane localization and unique regions) was coexpressed in insect cells with full-length Csk using differentially regulated promoters, resulting in almost complete phosphorylation of the C-terminal tail of Hck at Tyr 522 (Tyr 527 in c-Src), with little or no autophosphorylation at Tyr 411 (Tyr 416 in c-Src). Crystals of Hck were readily obtained, but they diffract X-rays relatively weakly. The quality of the X-ray diffraction was improved by adding the tyrosine kinase inhibitor quercetin (3,3',4',5,7-pentahydroxyflavone)⁹ or the ATP analogue AMP-PNP (β , γ -imidoadenosine-5'-triphosphate) to protein solutions during crystallization. The diffraction patterns are strongly anisotropic, with data measured to 3.0 Å resolution and to 2.6/2.9 Å for the quercetin and AMP-PNP complexes, respectively (Table 1). The crystals are in space group $P2_12_12_1$, with two molecules of Hck in the crystallographic asymmetric unit (see Methods).

The structure of the quercetin complex was determined using a combination of multiwavelength anomalous diffraction (MAD) and multiple isomorphous replacement with anomalous scattering (MIRAS) (Table 1). Refinement of this structure resulted in a final *R*-value of 22.9% and a free *R*-value of 31.5% at 3.0 Å resolution (Table 1). Subsequently, a model for the AMP-PNP complex was refined to *R*-values of 23.7% and 31.2% (2.6 Å resolution; Table 1). Residue numbering corresponding to chicken c-Src is used throughout this paper (Figure 1), and the overall sequence identity between chicken c-Src and Hck is 60%, 60% and 70% in the SH3, SH2 and catalytic domains.

Overall description of the structure

The SH2 and SH3 domains of Hck are bound intramolecularly to the N- and C-terminal extensions of the catalytic domain, on the side opposite to the substrate binding site of the kinase (Fig. 2). The structures of the component domains are closely similar to the previously determined structures of the corresponding regions from other proteins^{10–14}. The catalytic domain consists of an N-terminal lobe (N-lobe, residues 267 to 342), which contains the glycine-rich segment that interacts with the phosphate groups of the nucleotide, and a C-terminal lobe (C-lobe, residues 343 to 519) that provides the substrate-binding sites. Two adjacent regions in the C-terminal lobe are critical for catalysis and for regulation. The catalytic segment and strands β 7 and β 8 (residues 384 to 402) contain a conserved set of residues that are important for binding metal ions and the phosphate groups of ATP (reviewed in ref. 15). Tyr 416 in the Src-family kinases, which is autophosphorylated upon activation, is located in the activation segment (residues 403 to 430), and is not phosphorylated in the crystal structure. Residues 412 to 421 in the activation segment are disordered. The structures of the quercetin and AMP-PNP complexes are very similar, with the quercetin bound to the binding site for the adenine ring of AMP-PNP (Fig. 3).

The SH2 and SH3 domains are docked onto the catalytic domain by interactions that bear a striking resemblance to the well known modes of peptide target recognition by these domains (Fig. 4). The SH2 domain is bound to the C-terminal tail and is distant from the active site (Figs 2a and 4a–c), with the phosphotyrosine residue of the tail located more than 40 Å from the catalytic residues. The SH2 domain does not block the active site or the substrate-binding groove (Fig. 2b), and the structure indicates that its effect on catalysis is indirect.

An unanticipated feature of the architecture of Hck is the formation of a left-handed polyproline type II (PPII) helix¹⁶ by the SH2-kinase connector segment. Residues Pro 244 to Trp 254 are in the PPII conformation (Fig. 4d), with a break at Ser 247, which adopts a more extended conformation. SH3 domains recognize peptide targets that adopt a PPII conformation^{17,18}, and in a remarkable economy of binding modality the SH3 domain of Hck is docked onto the PPII helix formed by the SH2-kinase connector (Fig. 4d). There are very limited contacts between the SH2 and SH3 domains (Fig. 2a, b), and the relationship between the SH3 and SH2 domains in the Hck structure is quite distinct from that in the crystal structure of an SH3-SH2 fragment of Lck¹⁹.

The structure of the cyclic-AMP-dependent protein kinase A (PKA) complexed with ATP²⁰ provides a reference for the conformation of a fully active protein kinase¹⁵. Comparison with PKA shows that the catalytic domain of Hck is essentially in an active conformation, with the N- and C-lobes in a relative orientation that is close to that seen in active PKA. In contrast to insulin receptor tyrosine kinase¹³, the nucleotide- and substrate-binding sites of Hck are unimpeded by the activation segment.

When compared to PKA, the N-terminal region of the activation segment in Hck, is displaced upwards relative to the C-lobe, into a position where it overlaps with the position of helix α C in PKA. Helix α C is displaced in Hck as a result, with consequences for the catalytic state of the enzyme that are discussed below. The initial

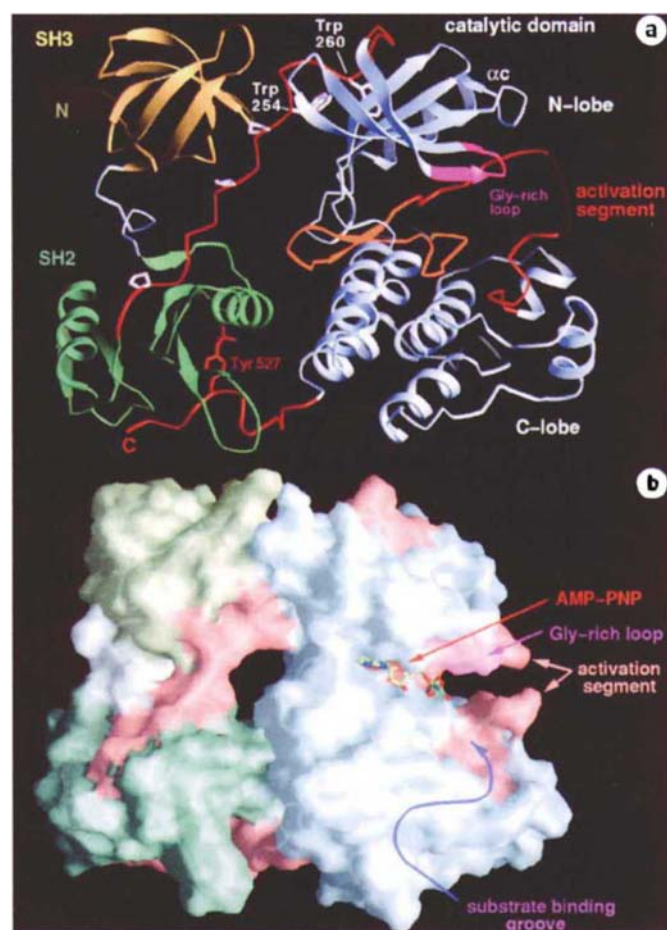


Figure 2 **a**, Ribbon representation of the Hck structure. All ribbon diagrams were generated using RIBBONS⁴⁴. **b**, Molecular surface of Hck, in the same orientation as **a**. The AMP-PNP molecule is shown as a stick figure. The surface area buried at the SH3 to SH2-kinase connector, kinase to SH2-kinase connector, SH2 to kinase tail and SH2 to kinase interfaces are 1,186 Å², 705 Å², 531 Å² and 1,091 Å², respectively. All molecular surfaces were generated using GRASP⁴⁵.

direction of the activation segment that causes the displacement of helix αC is similar to that in inactive cyclin-dependent protein kinase (CDK)²¹, where a similar displacement of helix αC is observed.

SH2 Interface

The interaction of the C-terminal tail of Hck and the SH2 domain resembles the 'two-pronged plug into two-holed socket' interaction of the Src and Lck SH2 domains and high-affinity peptides containing phospho-Tyr-Glu-Glu-Ile (pYEEI) motifs²²⁻²⁴. In the Hck tail interaction, the two prongs are formed by the side chains of the phosphotyrosine (residue 527) and the C-terminal Pro 531 (Fig. 4a). However, the intervention of three rather than two residues between Pro 531 and the phosphotyrosine causes the peptide to bulge up from the SH2 surface. The lack of a snug fit between the tail region and the SH2 domain is consistent with the fact that phosphopeptides with sequences corresponding to the tails of the Src family kinases are not the highest-affinity ligands for the SH2 domains of the kinases²⁵.

In addition to the interactions with the phosphorylated tail, the orientation of the SH2 domain is maintained by a number of electrostatic interactions with the C-lobe of the catalytic domain (Fig. 4c), mainly between helices αA and αE in the SH2 and catalytic domains, respectively. The charged residues at this interface are conserved in the Src-family kinases. v-Src, which has lost the regulatory action of the SH2 domain, has hydrophobic residues at two key positions within this region (residues 517 and 518, Fig. 1).

The SH3 interface

The PPII helix formed by the SH2-kinase connector is bound by the hydrophobic surface of the SH3 domain, and is flanked by the so-called 'n-Src' and 'RT-Src' loops (Fig. 4d). Pro 250 and Pro 253 of the SH2-kinase connector are in the P_{-1} and P_{+2} positions of the SH3 interface (Fig. 4d; see ref. 26 for the notation used). As for the SH2 domain, the interactions resemble that of SH3 domains with peptide ligands, but certain aspects of the interface appear to be less than optimal. Lysine side chains at the P_0 and P_{+3} positions provide some measure of hydrophobic interactions with the SH3 surface, and can potentially interact with Asp 91 and Glu 93 of the SH3 domain. They do not, however, take maximal advantage of the SH3 binding site because they do not point into the hydrophobic core of the SH3 domain. Likewise, the usual interaction between a basic residue at the P_{-3} position of the polyproline helix and Asp 99 in the SH3 domain is absent.

Gln 251 and Trp 254 in the connector occupy the P_{+1} and P_{+4}

positions (often proline in peptide ligands) and they interact with residues in the N-lobe of the kinase domain. Interactions made by Trp 254 appear to be particularly important, because it packs into a hydrophobic pocket framed by the β -sheet of the N-lobe and by the continuation of the SH2-kinase linker. The side-chain nitrogen of Trp 254 forms a hydrogen bond with the carbonyl group that immediately precedes Trp 260, a conserved residue that is implicated in the maintenance of the inactive conformation of the helix αC .

The sequence of the SH2-kinase linker is somewhat divergent between members of the Src family (Figure 1), as is the target specificity of the SH3 domains. For example, the HIV-1 protein Nef is a high-affinity ligand for the Hck SH3 domain ($K_d = 0.2 \mu M$), but binds only weakly or not at all to certain other Src-family SH3 domains^{27,28}. Swapping the SH3 domains between members of the Src family can lead to the loss of regulation²⁹. Only the first proline of the Pro-x-x-Pro motif in Hck is strictly conserved, and differences in the linker are presumably correlated with differences in the SH3 domains in other Src kinases.

Inhibition of catalytic activity

Protein kinases exhibit considerable flexibility in the relative orientations of the N- and C-lobes of the catalytic domain, with a 'closed' configuration being associated with the active form of the kinase because it brings the terminal phosphate group of ATP into proper alignment with the substrate acceptor group⁵. Inactive forms of protein kinases have been observed to be in various 'open' configurations in which the two lobes are swung apart with respect to their relative orientation in PKA.

The two lobes of the catalytic domain of Hck are essentially in a closed and active conformation. The net rotation that is required to superimpose the β -strands of the N-lobes of PKA and Hck after first aligning the C-lobes is small (about 6° and 9° , for the quercetin and AMP-PNP complexes, respectively, with a translational component $<1 \text{ \AA}$ in each case). The binding mode of AMP-PNP in Hck is generally similar to that of ATP in PKA, although the phosphate groups of the nucleotide are displaced outwards in Hck. This difference may be due to the presence of Ca^{2+} rather than Mg^{2+} at the metal-binding site as a consequence of the crystallization buffer.

The major structural difference between PKA and Hck at the active site involves the activation segment and the position of Glu 310 (Fig. 5a, b). In PKA, Glu 91 is equivalent to Glu 310 of Hck, and is completely buried in the structure. Glu 91 forms a salt bridge with Lys 72 (Lys 295 in Src), which coordinates the α - and β -phosphates of ATP in PKA²⁰. Although Glu 91 does not participate

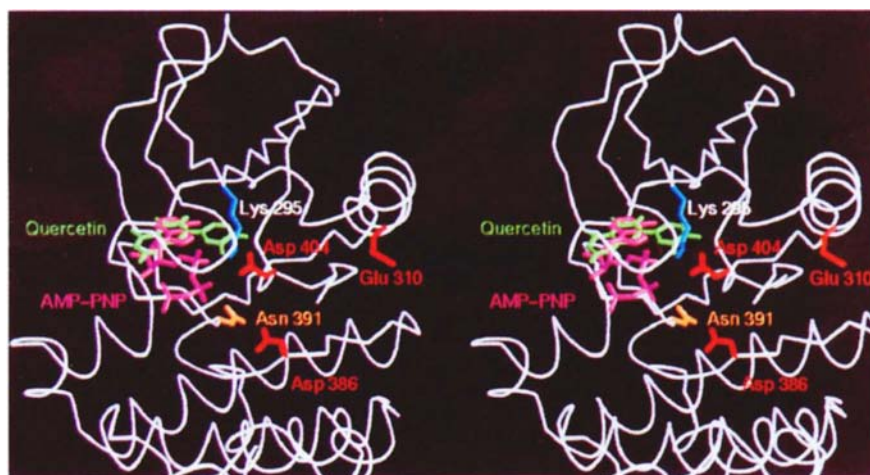
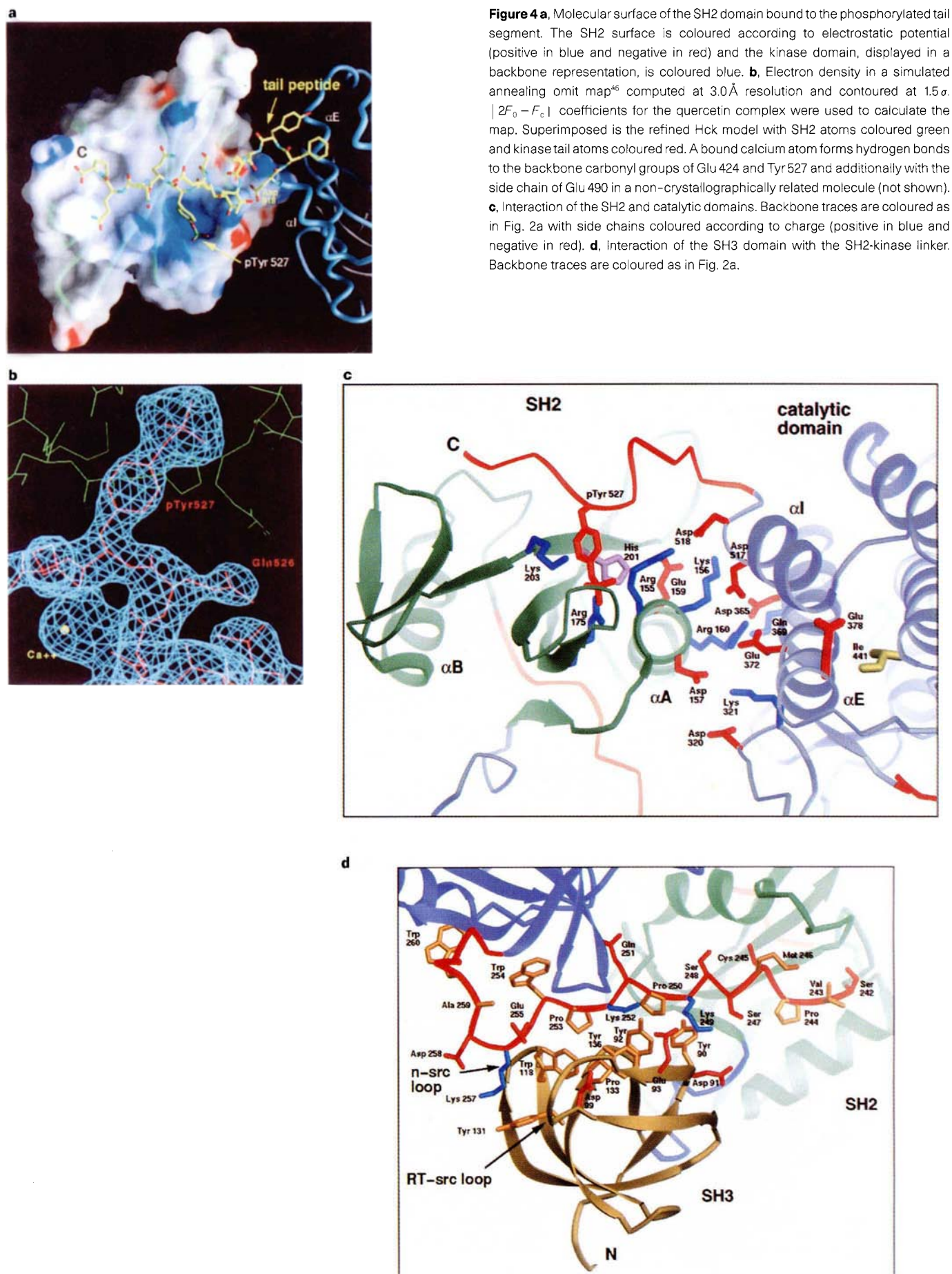


Figure 3 Stereo figure of the active site region of Hck with bound AMP-PNP and quercetin inhibitors. The backbone trace of the catalytic domain is shown in white.



directly in the catalytic mechanism, the positioning of the charged glutamate next to a lysine chain that coordinates ATP is likely to affect enzymatic activity profoundly. In Hck, because of the change in the orientation of helix α C, the side chain of Glu 310 is swung out of the active site and into solvent, such that the distance of its carboxyl group from the amino group of Lys 295 is 14.4 Å (the corresponding distance in PKA is 3.6 Å).

Comparison of Hck with the active form of the catalytic domain of Lck¹⁴ suggests that the conformation of helix α C is coupled to the phosphorylation state of Tyr 416 in the activation segment and to the conformation of the SH3-bound SH2-kinase connector (Fig. 5a, b). The activation segment and the SH2-kinase connector interact with the N-terminal and C-terminal ends of helix α C, respectively. In the outer conformation, Glu 310 is in a position to interact favourably with Arg 385 and Tyr 382 in the activation

segment. Upon phosphorylation of Tyr 416, Arg 385 interacts with it instead, and the activation segment no longer blocks the inward rotation of helix α C (ref. 14). Trp260, which is positioned by the SH2-kinase connector points into the hydrophobic core of the N-lobe and displaces Leu 325 and Met 314, in helix α C (Fig. 5c). This conformation of Trp 260 is inconsistent with the active and inner conformation of helix α C, and is likely to be destabilized by removal of the SH3 domain and melting of the PPII helix. Conversely, phosphorylation of Tyr 416 would favour the inner conformation of α C, with destabilization of this conformation of Trp 260 and release of the SH3 domain. This is consistent with the results of proteolytic cleavage of c-Src by trypsin, which show that Trp 260 is close to a major site of proteolytic cleavage in activated forms of c-Src, whereas this region is protected in the repressed form of the protein³⁰.

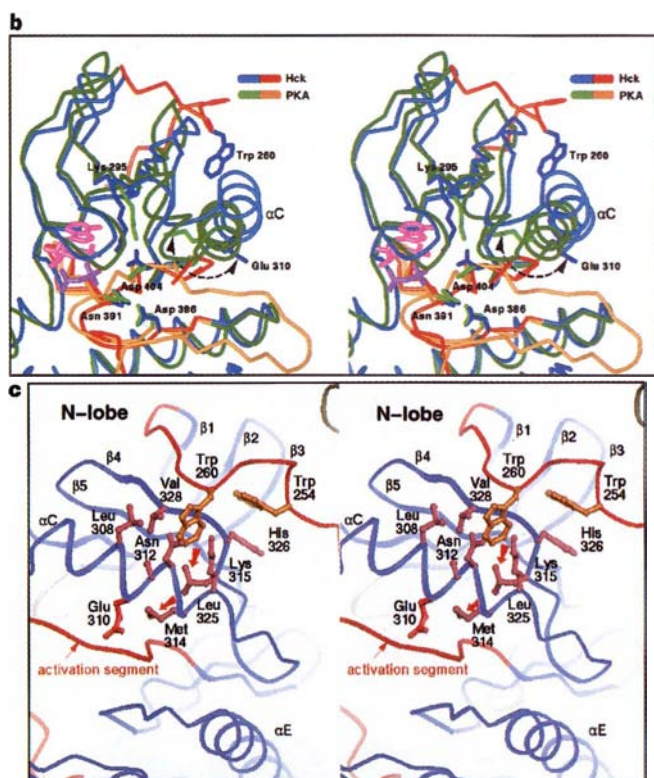
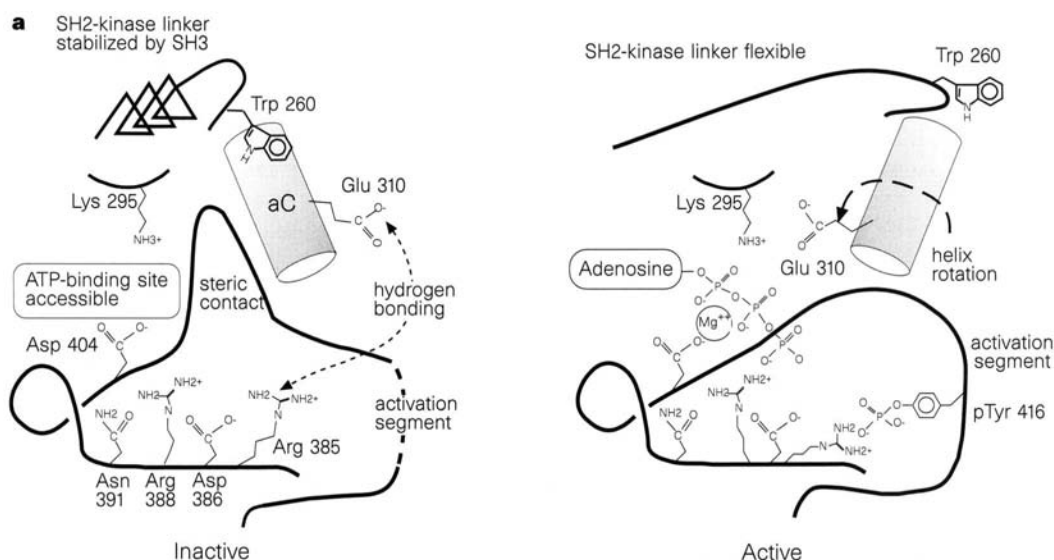


Figure 5 Regulation of catalytic activity. **a**, Schematic diagram indicating the difference between the inactive (left) and active (right) states of the Src-family tyrosine kinases. The diagram for the inactive configuration is based on the structures of PKA²⁰ and Lck¹⁴. **b**, Superposition of the kinase domains of Hck and PKA²⁰. Hck is shown in blue with the activation segment and the SH2/kinase connector region shown in red and the AMP-PNP inhibitor molecule shown in pink. PKA is shown in green with the activation segment shown in orange. The arrow indicates the movement of Glu 310. The N- and C-lobes were both used in the superpositioning. **c**, Interaction of the helix α C with the SH2-kinase connector and the kinase activation segment. The displacements, relative to the PKA structure, of the side chains of Leu 325 and Met 314 by Trp 260 are indicated by red arrows.

The orientation of helix α C and the location of Glu310 in Hck is very similar to that in the inactive form of CDK^{21,31}. Why is the outer, inactive, conformation of helix α C stable in CDK without the need for the additional interactions provided by the SH3 domain in Hck? In CDK, the formation of an α -helix in the activation segment results in the insertion into the N-lobe of two bulky side chains. Leu 148 and Phe 152 (ref. 21). In contrast, the activation segment in Hck does not form a helix, and consequently there is less steric hindrance. Thus, whereas CDK relies primarily on the unphosphorylated activation segment for α C displacement, Hck appears to act on the helix at both ends.

Consequences for activation of kinases

The SH2-tail interaction is clearly important for repression of Src activity, because regulation is lost upon dephosphorylation, mutation or displacement of the phosphotyrosine^{32–37}. The structure makes clear that the SH2 domain does not block or alter the catalytic domain. The accompanying paper³⁸ shows that the HIV-1 protein Nef, a high-affinity ligand for the Hck SH3 domain, is a potent activator of the catalytic activity of Hck and is able to further activate Hck that is dephosphorylated on Tyr 527. Mutations in the c-Src SH3 domain that activate the kinase map to the SH3:peptide interface, a result that is completely consistent with the Hck structure²⁹. Furthermore, the RT-src and n-Src loops that flank the peptide-binding region of the SH3 domain have been shown to be important for regulation²⁹, and in the Hck structure these loops make contacts with the N-lobe of the kinase.

Autophosphorylation of Hck on Tyr 416 is very slow and proceeds intermolecularly when the SH2 and SH3 domains are in place³⁸. Displacement of the SH3 domain by Nef greatly accelerates autophosphorylation at Tyr 416 and, in addition, increases the catalytic activity of the autophosphorylated form by roughly six-fold³⁸. Removal of the SH3 domain is likely to destabilize the polyproline helix, as will disruption of the SH2-tail interaction. This would favour the conversion of the active site to a catalytically competent form. The apparent flexibility of the activation segment in the vicinity of Tyr 416 suggests that autophosphorylation could occur intramolecularly once the regulatory domains are removed.

The activation of Src-family tyrosine kinases by displacement of SH3 domains may be a general feature of their regulation. The Ras GTPase-activating protein, GAP, is an Hck substrate, and also binds to the SH3 domain of Hck³⁹. Although the effect of GAP on Hck activity was not tested, the results with Nef suggest that GAP binding would activate Hck. Using a transcriptional readout assay, the ability of a Src-SH3/SH2-binding protein to activate c-Src has been demonstrated⁴⁰.

The Hck structure also provides an explanation for the effects of two mutations in a c-Src that are not in the regulatory apparatus or at the active site, but yet lead to activation³⁷. Glu 378 is located in the α E helix, and its side chain is positioned directly at the base of the α I helix, where the negative charge would interact favourably with the helix dipole (Fig. 4c). Mutation of this residue to glycine activates c-Src, perhaps by destabilization of helix α I or alteration of its direction, thereby affecting the SH2 interaction. Likewise, the side chain of Ile 441 is located near that of Glu 378. Mutation of this residue to Phe is also activating, and is likely to disturb the interaction of Glu 378 with helix α I.

Discussion

The structure of Hck reveals how the Src-family tyrosine kinases regulate their catalytic activity by relying on localized interactions that build on the intrinsic binding properties of the SH2 and SH3 domains. An interesting aspect of the internal engagement of the SH2 and SH3 domains is that their binding sites do not appear to be quite optimal in structural terms. This is likely to have important consequences for function, in that it allows for appropriate com-

petition by external ligands that are biologically important targets for these domains.

A fascinating aspect of the regulation of Hck is the similarity in the mechanism of catalytic control with that seen in the cyclin-dependent protein kinases. However, these two centrally important protein kinases use quite distinct molecular interactions to bring about the appropriate response in the switching helix. The structure of Hck underscores the versatility with which diverse regulatory mechanisms are brought to bear on the conserved core of the protein kinases. □

Methods

Protein expression. Sf9 cells were co-infected with separate recombinant baculoviruses expressing human Hck from the polyhedrin promoter and chicken Csk under the control of the basic protein promoter. Cells were grown in suspension to a density of not more than 2×10^6 cells per ml and co-infected with the individual viruses at a multiplicity of 10 each. After further growth for 40–48 h, cells were collected by centrifugation and stored at -80°C . For protein purification about 10^{10} cells were lysed in 100 ml buffer A (20 mM Tris 8.5, 10% glycerol, 1 mM DTT, 1 mM EDTA, 0.1 mM PMSE, 0.1 mM sodium orthovanadate) by sonication and the extract was clarified by high-speed centrifugation. This lysate was diluted to 500 ml using buffer A, loaded on a 75 ml macroprep high Q column (Biorad) and eluted by a linear KCl gradient. Fractions containing Hck were pooled, brought to a final concentration of 1.5 M ammonium sulphate, loaded onto a phenyl sepharose HP column (10 ml, Pharmacia) and eluted by a decreasing salt gradient. The peak fraction was concentrated and passed over a Superdex75 gel filtration column (Pharmacia) equilibrated in 10 mM HEPES, pH 7.5, 1 mM DTT. The peak fraction was concentrated to 50 mg ml^{-1} Hck and used for crystallization or biochemical studies. About 10 mg purified Hck was obtained from 11 cell culture. Western blot analysis demonstrated that the purified protein was phosphorylated on tyrosine. Mass spectrometric analysis of proteolysed purified protein showed that all peptides derived from the activation segment, around Tyr 416, were only detected in their unphosphorylated form.

Crystallization, data collection and processing. Hanging drops (1 μl) of 50 mg ml^{-1} protein and 10 mM quercetin or 2AMP-PMP were mixed with equal volumes of reservoir buffer containing 150 mM calcium acetate, 100 mM cacodylate (pH 6.5), 7% w/v PEG 8000, and 16% v/v ethylene glycol. Single crystals of rough dimensions $0.05 \times 0.2 \times 0.2 \text{ mm}^3$ were routinely obtained within 4 days after streak seeding using smaller crystals obtained initially. The crystals contain two molecules of Hck per asymmetric unit, and are in space group $P2_12_12_1$ ($a = 73.5/72.8 \text{ \AA}$, $b = 93.3/92.4 \text{ \AA}$, $c = 176.9/178.3 \text{ \AA}$ for quercetin/AMP-PNP complexes). Crystals were cryo-protected in reservoir buffer enriched to 20% w/v PEG 8000, 20% v/v ethylene glycol and 2 mM quercetin or AMP-PNP before flash freezing. Heavy-atom derivatives were prepared by soaking crystals for 15 h in 5 mM heavy-atom/cryo-protection solutions. The diffraction patterns exhibit significant anisotropy; diffraction to $2.5 \text{ \AA}$ was apparent along c^* but diffraction was limited to $2.9 \text{ \AA}$ along other dimensions for native crystals. Data were collected for a quercetin complex, AMP-PMP complex and the ethyl mercury phosphate, thimerosal and mersalyl acid derivatives using an R-AXIS IIC imaging plate detector system with RU200 rotating anode generator (Table 1). Data for the quercetin complex were also collected at Brookhaven National Laboratories (BNL) on beam line X25 ($\lambda = 1.000 \text{ \AA}$) using a MAR imaging plate detector system. A MAD experiment on a thimerosal derivatized crystal was performed at BNL, on beamline X4A ($\lambda_1 = 1.0063 \text{ \AA}$, $\lambda_2 = 1.0093 \text{ \AA}$ and $\lambda_3 = 0.992 \text{ \AA}$) with data collected using Fuji image plates. Data processing and reduction was carried out with the HKL, DENZO, and SCALEPACK programs (Z. Otwinowski and W. Minor, unpublished programs). MIRAS and MAD phases were calculated using the PHASES package⁴¹. In the MAD experiment, data quality was greatly impaired by crystal decay, with the effective resolution decreasing from $3.0 \text{ \AA}$ to $3.6 \text{ \AA}$ over the course of the experiment. As a result the λ_3 data set could not be used. MIRAS and MAD phases were combined to obtain the best electron density map.

Model building and refinement. Model building was performed using O⁴², using known SH3, SH2 and catalytic domain structures as guides. Bulk solvent correction and anisotropic B -factor scaling was applied during refinement using X-PLOR⁴³. The elements of the final anisotropic B -factor tensor applied

to data for the quercetin and AMP-PNP complexes are $(4.6 \text{ \AA}^2, -23.7 \text{ \AA}^2, 22.2 \text{ \AA}^2)$ and $(2.2 \text{ \AA}^2, -18.3 \text{ \AA}^2, 17.4 \text{ \AA}^2)$ along a^* , b^* and c^* , respectively. Non-crystallographic symmetry (NCS) restraints were applied and the r.m.s. deviation between C α atoms related by NCS is 0.34–0.40 Å for the various domains, which were treated separately. The r.m.s. deviation between C α atoms in the quercetin and AMP-PNP complexes is 0.45 Å. The final refinement statistics are shown in Table 1. The four N-terminal residues, 78 to 81, and residues 412 to 421 of the kinase activation region have not been modelled. Residues with disordered side chains modelled as Ala include D192, R194, Q915, N209, E225, N234, Q277, R409, V410, I411, K423, E476. No amino-acid residues occupy disallowed regions of the Ramachandran plot and 83% occupy the most favoured regions.

Received 11 December 1996; accepted 22 January 1997.

1. Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of Src. *Biochim. Biophys. Acta* **1287**, 121–149 (1996).
2. Quintrell, N. et al. Identification of a human gene (*HCK*) that encodes a protein-tyrosine kinase and is expressed in hemopoietic cells. *Mol. Cell. Biol.* **7**, 2267–2275 (1987).
3. Zeigler, S. E., March, J. D., Lewis, D. B. & Perlmutter, R. M. Novel protein-tyrosine kinase gene (*hck*) preferentially expressed in cells of hematopoietic origin. *Mol. Cell. Biol.* **7**, 2276–2285 (1987).
4. Lowell, C. A., Soriano, P. & Varmus, H. E. Functional overlap in the src gene family: interaction of *hck* and *lck* impairs natural immunity. *Genes Dev.* **8**, 387–398 (1994).
5. Lowell, C. A., Niwa, M., Soriano, P. & Varmus, H. E. Deficiency of the Hck and Src Tyrosine Kinases results in extreme levels of extramedullary hematopoiesis. *Blood* **87**, 1780–1792 (1996).
6. Pawson, R. Protein modules and signalling networks. *Nature* **373**, 573–580 (1995).
7. Cohen, G. B., Ren, R. & Baltimore, D. Modular binding domains in signal transduction proteins. *Cell* **80**, 237–248 (1995).
8. Kuriyan, J. & Cowburn, D. Modular Peptide Binding Domains. *Annu. Rev. Biophys. Biomol. Struct.* (in the press).
9. Cunningham, B. D., Threadgill, M. A., Groundwater, P. W., Dale, I. L. & Hickman, J. A. Synthesis and biological evaluation of a series of flavones designed as inhibitors of protein tyrosine kinases. *Anticancer Drug Design* **7**, 365–384 (1992).
10. Waksman, F. et al. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature* **358**, 646–653 (1992).
11. Musacchio, A., Noble, M., Paupit, R., Wierenga, R. & Saraste, M. Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359**, 851–855 (1992).
12. Knighton, D. R. et al. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 407–414 (1991).
13. Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746–754 (1994).
14. Yamaguchi, H. & Hendrickson, W. A. Structural basis for activation of the human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* **384**, 484–489 (1996).
15. Johnson, L. N., Noble, M. E. M. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149–158 (1996).
16. Adzhubei, A. A. & Sternberg, M. J. E. Left-handed polyproline II helices commonly occur in globular proteins. *J. Mol. Biol.* **229**, 472–493 (1993).
17. Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W. & Schreiber, S. L. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933–945 (1994).
18. Musacchio, A., Saraste, M. & Wilmanns, M. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nature Struct. Biol.* **1**, 546–551 (1994).
19. Eck, M. J., Atwell, S. K., Shoelson, S. E. & Harrison, S. C. Crystal structure of the regulatory domains of the Src-family tyrosine kinase lck. *Nature* **368**, 764–769 (1994).
20. Zheng, J. et al. 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr.* **D49**, 362–365 (1993).
21. DeBondi, H. L. et al. Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595–602 (1993).
22. Waksman, G., Shoelson, S. E., Plant, N., Cowburn, D. & Kuriyan, J. Binding of a high affinity phosphotyrosyl peptide to the src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell* **72**, 779–790 (1993).

23. Eck, M. J., Shoelson, S. E. & Harrison, S. C. Recognition of a high affinity phosphotyrosyl peptide by the Src homology 2 domain of p56lck. *Nature* **362**, 87–91 (1993).
24. Songyang, Z. et al. SH2 domains recognize specific phosphopeptide sequences. *Cell* **72**, 767–778 (1993).
25. Ladbury, J. E. et al. Measurement of the binding tyrosyl phosphopeptides to SH2 domains: a reappraisal. *Proc. Natl Acad. Sci. USA* **92**, 3199–3203 (1995).
26. Lim, W. A., Richards, D. M. & Fox, R. O. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372**, 375–379 (1994).
27. Lee, C.-H. et al. A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. *EMBO J.* **14**, 5006–5015 (1995).
28. Lee, C.-H., Saksela, K., Mirza, U. A., Chait, B. T. & Kuriyan, J. Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell* **85**, 931–942 (1996).
29. Erpel, T. A., Superti-Furga, G. & Courtneidge, S. A. Mutational analysis of the Src SH3 domain: the same residues of the ligand binding surface are important for intra- and inter-molecular interactions. *EMBO J.* **14**, 963–975 (1995).
30. MacAuley, A. & Cooper, J. A. Structural differences between repressed and derepressed forms of p60c-src. *Mol. Cell. Biol.* **9**, 2448–2456 (1989).
31. Juffrey, P. D. et al. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376**, 294–295 (1995).
32. Courtneidge, S. A. Activation of the pp60c-src kinase by middle T antigen binding or by dephosphorylation. *EMBO J.* **4**, 1471–1477 (1985).
33. Kmiecik, T. E. & Shalloway, D. Activation and suppression of pp60c-src transforming ability by mutation of its primary sites of tyrosine phosphorylation. *Cell* **49**, 65–73 (1987).
34. Piwnicka-Worms, H., Saunders, K. B., Roberts, T. M., Smith, A. E. & Cheng, S. H. Tyrosine phosphorylation regulates the biochemical and biological properties of pp60c-src. *Cell* **49**, 75–82 (1987).
35. Cartwright, C. A., Eckhart, W., Simon, S. & Kaplan, P. L. Cell transformation by pp60c-Src mutated in the carboxy-terminal regulatory domain. *Cell* **49**, 83–91 (1987).
36. Cooper, J. A. & King, C. S. Dephosphorylation or antibody binding to the carboxy terminus stimulates pp60c-src. *Mol. Cell. Biol.* **6**, 4467–4477 (1986).
37. Parsons, T. J. & Weber, M. J. Genetics of src: structure and functional organization of a protein tyrosine kinase. *Curr. Top. Microb. Immunol.* **147**, 80–127 (1989).
38. Moore, I. et al. Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. *Nature* **385**, 650–653 (1997).
39. Briggs, S. D., Bryant, S. S., Jove, R., Sanderson, S. D. & Smithgall, T. E. The Ras GTPase-activating protein (GAP) is an SH3 domain-binding protein and substrate for the Src-related tyrosine kinase, Hck. *J. Biol. Chem.* **270**, 14718–14724 (1995).
40. Alexandropoulos, K. & Baltimore, D. Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel p130Cas-related protein. *Science* **270**, 1341–1355 (1996).
41. Furey, W. & Swaminathan, S. in *American Crystallographic Association Meeting Abstracts* **73** (1990).
42. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**, 110–119 (1991).
43. Brünger, A. R. X-PLOR (Yale University, New Haven, CT, 1992).
44. Carson, M., Ribbons, D. O. *J. Appl. Crystallogr.* **24**, 958–961 (1991).
45. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Protein Struct. Funct. Genet.* **11**, 281–296 (1991).
46. Brünger, A. T. Crystallographic refinement by simulated annealing: application to a 2.8 Å resolution structure of aspartate aminotransferase. *J. Mol. Biol.* **203**, 803–816 (1988).

Acknowledgements. We thank H. Viguet for technical assistance with cell culture; M. Huse for crystals of the AMP-PNP complex; S. Jacques for protein purification; J. Bonanno for X-ray data collection; B. Chait and W. Zhang for mass spectrometric analysis; K. Saksela for the Hck clone; C. Ogata and L. Berman for assistance with synchrotron data collection; D. Yang for computational assistance; and S. Burley and D. Cowburn for useful suggestions. Advice provided by members of the Kuriyan Laboratory was invaluable at all stages. We thank H. Yamaguchi and W. Hendrickson for coordinates of Lck before publication, and for comments regarding tyrosine kinase regulation. We thank M. Eck and S. Harrison for exchanging Co coordinates and for discussions. E.S. is a Fellow of the Human Frontiers Science Program. J.K. thanks H. Hanafusa for introducing him to this field and for suggesting this line of inquiry.

Correspondence and requests for materials should be addressed to J.K. (e-mail: kuriyan@rockvax.rockefeller.edu). Coordinates are being deposited in the protein databank and are available from <http://www.rockefeller.edu/kuriyan>.

YOURS TO HAVE AND TO HOLD BUT NOT TO COPY

The publication you are reading is protected by copyright law.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

A massive black hole at the centre of the quiescent galaxy M32

Roeland P. van der Marel*, P. Tim de Zeeuw†, Hans-Walter Rix‡ & Gerald D. Quinlan§

* Institute for Advanced Study, Olden Lane, Princeton, New Jersey 08540, USA

† Sterrewacht Leiden, Postbus 9513, 2300 RA Leiden, The Netherlands

‡ Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

§ Department of Physics and Astronomy, Rutgers University, PO Box 849, Piscataway, New Jersey 08855, USA

Massive black holes are thought to reside at the centres of many galaxies^{1,2}, where they power quasars and active galactic nuclei. But most galaxies are quiescent, indicating that any central massive black hole present will be starved of fuel and therefore detectable only through its gravitational influence on the motions of the surrounding stars. M32 is a nearby, quiescent elliptical galaxy in which the presence of a massive black hole has been suspected^{3–9}; however, the limited resolution of the observational data and the restricted classes of models used to interpret this data have made it difficult to rule out alternative explanations, such as models with an anisotropic stellar velocity distribution and no dark mass or models with a central concentration of dark objects (for example, stellar remnants or brown dwarfs). Here we present space-based high-resolution optical spectra of M32, which show that the stellar velocities near the centre of this galaxy exceed those inferred from previous ground-based observations. We use a range of general dynamical models to determine a central dark mass concentration of $(3.4 \pm 1.6) \times 10^6$ solar masses, contained within a region only 0.3 pc across. This leaves a massive black hole as the most plausible explanation of the data, thereby strengthening the view that such black holes exist even in quiescent galaxies.

We observed M32 with the Faint Object Spectrograph (FOS) on the Hubble Space Telescope (HST), on 22 August 1995. The COSTAR optics corrected the spherical aberration of the HST primary mirror. Spectra of the wavelength range 4,572–6,821 Å were taken through square apertures of 0.086 and 0.215 arcsec, aligned on the nucleus and along the photometric major axis. The kinematical analysis was done by fitting each galaxy spectrum with the convolution of a template spectrum and a gaussian line-of-sight velocity profile. The details of the observations will be discussed by R.P.v.d.M., P.T.d.Z. and H.-W.R. (manuscript in preparation). The inferred kinematics are compared with the best ground-based data^{9,10} in Fig. 1. The rotation curve derived from the HST spectra rises more steeply than that measured from the ground. The velocity dispersion from the HST aperture closest to the centre is $156 \pm 10 \text{ km s}^{-1}$. The average of the four independent dispersion measurements inside the central 0.1 arcsec is 126 km s^{-1} , exceeding significantly the $85\text{--}95 \text{ km s}^{-1}$ measured from the ground.

Dynamical models are needed to infer the mass distribution from the kinematical data, and to test whether the observed velocities are consistent with a keplerian rise around a black hole. To this end we developed a new technique to construct fully general axisymmetric models that fit a given set of observations (H.-W.R. *et al.*, R.P.v.d.M. *et al.*, N. Cretton *et al.*, manuscripts in preparation). It is based on Schwarzschild's approach¹¹, and can be viewed as an axisymmetric generalization of the spherical technique of Richstone and collaborators^{3,6,12}. It consists of four steps: (1) an axisymmetric mass density is chosen that fits the M32 HST photometry¹³ after projection; (2) the gravitational potential is calculated, including a nuclear dark mass; (3) a representative library of orbits is calculated;

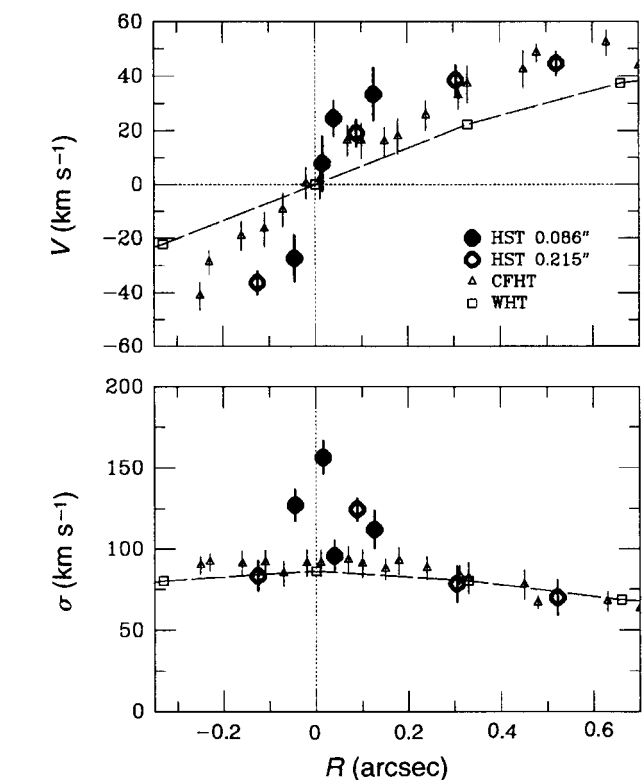


Figure 1 Stellar rotation velocities V (top panel) and velocity dispersions σ (bottom panel) in M32, as function of major-axis distance. The new HST data (aperture sizes indicated in the figure) are compared with ground-based data obtained at the William Herschel Telescope¹⁰ (WHT; resolution ~ 0.8 arcsec; connected by a dashed line for illustration; error bars smaller than plot symbols) and the Canada-France Hawaii Telescope⁹ (CFHT; resolution ~ 0.5 arcsec). The HST data were obtained with the G570H grating on the FOS red detector, and were calibrated using arc lamp spectra obtained during Earth occultations. The pointing accuracy was limited by target acquisition uncertainties (~ 0.02 arcsec) and thermal effects on the spacecraft and Fine Guidance Sensors (increasing to ~ 0.1 arcsec during the 14 h of the observations). The actual aperture position for each observation was calculated *post hoc* to ~ 0.01 arcsec accuracy from models based on HST photometry¹³ for the target acquisition data, the observed count rate in each spectrum, and a verification image obtained at the end of the observations. The template spectrum used for the kinematical analysis was a mix of ground-based spectra of stars of different spectral types, chosen to best match the M32 spectrum. The kinematical results were corrected for line-spread-function differences between the galaxy and template spectra. The line spread function for each galaxy spectrum was calculated from models for the distribution of the galaxy light within the aperture, taking into account the grating broadening function and the finite size of the detector diodes. The relatively low velocity dispersion inferred for the HST aperture second-closest to the nucleus might have been affected by a systematic error, but none could be identified.

and (4) a non-negative least-squares fit is performed to determine the combination of orbits that reproduces the assumed density and best fits the observational constraints.

Previous studies of M32 have used either axisymmetric models with distribution functions f that depend only on the two classical integrals of motion (the energy E and angular momentum L_z around the symmetry axis), $f(E, L_z)$, or spherical models with fully general distribution functions. Models of the first type take into account the flattening of M32 (projected axial ratio ~ 0.73), but have a special dynamical structure. They successfully fit the ground-based data when a nuclear dark mass is included^{7–9}, but they leave open the possibility that more general models might fit the data

without a dark mass. Models of the second type allow for general anisotropy, but ignore the flattening of M32. These spherical models were used only to interpret some of the older M32 data⁵, and could not fit these without a nuclear dark mass⁶. However, we found that our less-restricted axisymmetric models can easily fit these same data without a dark mass. So even though previous studies have shown that existing data are consistent with a nuclear dark mass in M32, it has not been shown that such a dark mass is required.

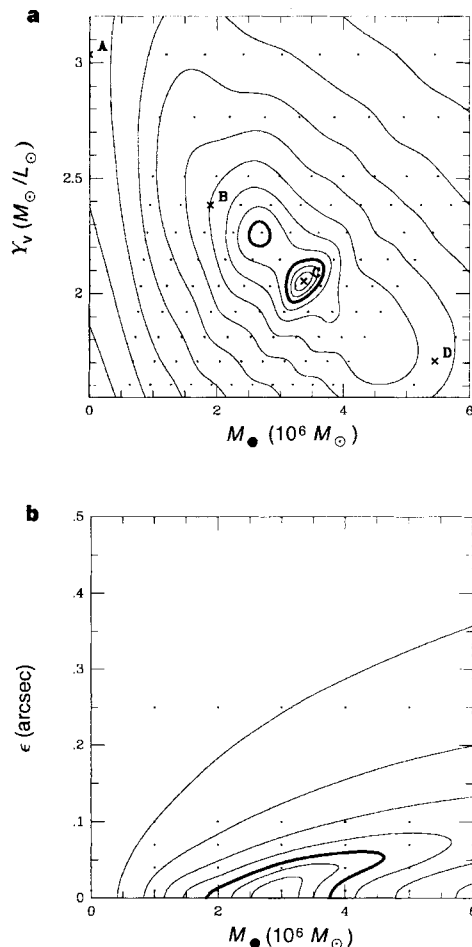


Figure 2 a. Contours of the χ^2 that measures the quality of the fit to the combined HST and ground-based kinematical M32 data, for edge-on models with a fully general dynamical structure. The model parameters $M_\bullet$ and T_V are the dark nuclear point mass and the V-band stellar mass-to-light ratio. The dots indicate models that were calculated. The predictions of the models labelled A–D are shown in Fig. 3. The contours were obtained through spline interpolation. The first three contours define the formal 68.3%, 95.4% and (heavy contours) 99.73% confidence regions. Subsequent contours are characterized by a factor two increase in $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$. The confidence values assume that there are only gaussian random errors in the data, and no numerical errors in the models. In reality there are numerical errors in the models, but they are known to be small (see Fig. 3 legend). Tests show that these errors can be responsible for the presence of the second (local) minimum in the χ^2 contours (although this might also be real), but that they cannot make the predictions of models with $M_\bullet < 1.8 \times 10^6 M_\odot$ or $M_\bullet > 5.0 \times 10^6 M_\odot$ consistent with the data at the formal 99.73% confidence level. We therefore conclude that $M_\bullet = (3.4 \pm 1.6) \times 10^6 M_\odot$. **b.** Contours of the χ^2 that measures the quality of the fit to the HST velocity dispersions, for edge-on $f(E, L_z)$ models with an extended dark nuclear object with scale length ϵ (and with $T_V = 2.51$ to best fit the data outside the central arcsec for this assumed dynamical structure). A point mass ($\epsilon = 0$) provides the best fit; models with $\epsilon \geq 0.06$ arcsec are ruled out.

To improve this, we used our general axisymmetric technique to determine which models can fit both the new HST data and the most recent ground-based data^{9,10}. The models take the observational characteristics into account, and include velocity-profile shape data (not available from the HST because of the limited spectral resolution of the FOS). Figure 2a shows a contour plot of the χ^2 of the fit for edge-on models with a dark central point mass. The free parameters are the V-band mass-to-light ratio T_V of the stellar population, and the dark mass $M_\bullet$. The models labelled A–D are the best-fitting models for four fixed values of $M_\bullet$. Figure 3 compares the predictions of these models with the observed major-axis kinematics for each of the data sets. Models without a dark mass are ruled out; even the best-fitting model without a dark mass, model A, fails to fit the nuclear HST velocity dispersions by $>40 \text{ km s}^{-1}$. Model C has $M_\bullet = 3.4 \times 10^6 M_\odot$, and provides the overall lowest χ^2 . However, we would not necessarily expect our algorithm to constrain a single, best-fit choice of the potential, as a three-dimensional distribution function might have the freedom to adjust to changes of the potential without affecting the fit to the data. Numerical experiments showed that the structure in the χ^2 plot near the two minima could in fact be unreliable. We therefore do not claim that the local minimum labelled C is necessarily

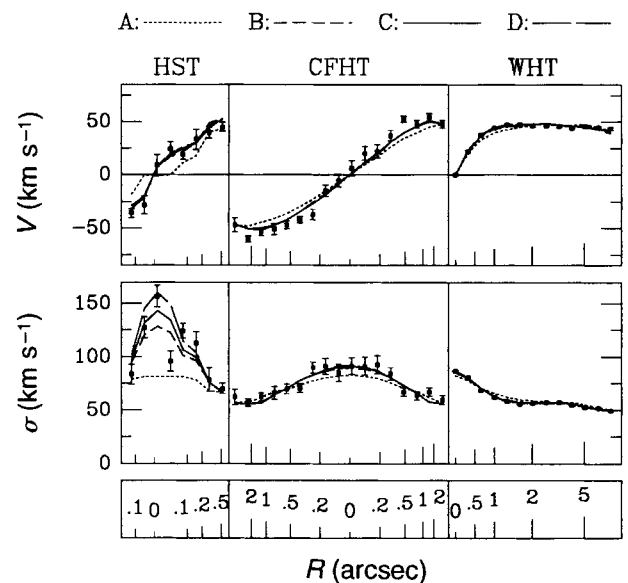


Figure 3 Predicted rotation velocities V and velocity dispersions σ of models A–D in Fig. 2, compared with the major-axis data from the HST, CFHT and WHT. Data points are plotted equidistantly along the abscissa. Corresponding major-axis distances are given in the bottom panel. The WHT measurements are averages of data obtained at positive and negative radii. The displayed quantities form only a subset of all the data that were fitted (for the ground-based observations^{9,10}, velocity-profile shape data and position angles other than the major axis are also available). Model C is the edge-on model with the lowest χ^2 , and has a dark nuclear point mass $M_\bullet = 3.4 \times 10^6 M_\odot$. Models B and D are (approximately) the best-fitting models for $M_\bullet = 1.9$ and $5.4 \times 10^6 M_\odot$, respectively. Model A is the best fit without a dark mass. This model is ruled out. It is already marginally ruled out by the ground-based CFHT and WHT data, but only the HST velocity dispersion measurements make the case clear-cut. The accuracy of the model predictions was assessed by having them reproduce the known results for $f(E, L_z)$ models^{7a}. From these tests we conclude that numerical errors in our models are sufficiently small not to influence the conclusions of this Letter (they are $<2 \text{ km s}^{-1}$ in the projected kinematics and <0.01 in the Gauss–Hermite moments¹⁰ of the velocity profiles).

special, but conservatively estimate the allowed range for $M_\bullet$ to be $(3.4 \pm 1.6) \times 10^6 M_\odot$. The uncertainty takes into account not only the observational errors in the data, but also the possible influence of small ($< 2 \text{ km s}^{-1}$ in the projected kinematics) numerical errors in the models (due to, for example, gridding and discretization). M32 need not be edge-on, but models constructed for inclination $i = 55^\circ$ gave similar results and the same allowed mass range. A full exploration of the range of possible inclinations is computationally prohibitive, but our results and previous work⁸ suggest that the inferred $M_\bullet$ depends only weakly on the assumed inclination. Allowing for possible triaxiality is not likely to change these conclusions; they are due mostly to the radial behaviour of the potential, and not to its shape. The best-fitting models are not $f(E, L_z)$ models, but are similar in that they are tangentially anisotropic and require a similar dark mass.

To constrain the size of the dark mass, models were 'built' with the point-mass potential replaced by a Plummer potential with the same mass, but with scale length ϵ . These were restricted to $f(E, L_z)$ models for simplicity; constraints on ϵ from the more general models should not differ by much. Figure 2b shows a contour plot of the χ^2 of the fit to the HST velocity dispersions, for edge-on models. The best fit has $\epsilon = 0$ (a point mass); models with $\epsilon \geq 0.06 \text{ arcsec}$ are ruled out.

The mass concentration thus has a half-mass radius $r_h = 1.30\epsilon \leq 0.08 \text{ arcsec}$, implying a central density $\geq 10^8 M_\odot \text{ pc}^{-3}$ ($1 \text{ arcsec} = 3.39 \text{ pc}$) and a mass-to-light ratio ≥ 20 inside 0.1 arcsec . It cannot be a cluster of ordinary stars; that would evolve rapidly by stellar collisions¹³, and a change in mass-to-light ratio from ~ 2 at 1 arcsec to ≥ 20 at 0.1 arcsec would yield large broad-band colour gradients, ruled out by observations^{14,15}. It must be dark, either a massive black hole or a cluster of dark objects (for example, stellar remnants, brown dwarfs, planets). No theory predicts the formation of such clusters—except perhaps^{16,17} for a cluster of small black holes of mass $m \sim (5\text{--}10)M_\odot$, each—and most would not last for the age of the galaxy. A dark cluster is not an acceptable alternative to a massive black hole if it will collapse to a black hole in a short time, or expand or eject its mass through explosions or evolve in some other way that would make it observable (for then we would have to assume that it formed recently). Goodman and Lee¹⁸ considered possible dark clusters for M32, and argued that the clusters required $r_h \geq 0.1 \text{ arcsec}$ to be acceptable; these were consistent with ground-based data, which allowed⁶ ϵ to be as large as 0.4 arcsec . We summarize and update the arguments here to show that all but the most implausible clusters are now ruled out.

A cluster of stellar remnants of mass $m \geq 1 M_\odot$ would undergo core collapse in a short time, $5.5 \times 10^8 \text{ yr}$ ($1 M_\odot/m$) ($r_h/0.08 \text{ arcsec}$)^{3/2} for a Plummer model (and shorter for more concentrated models¹⁹). Neutron stars or small black holes would form binaries and merge during the collapse by dissipating energy through gravitational radiation²⁰. The binaries would not be able to stop the collapse until they had grown large through multiple mergers, and even then they would at most cause rapid oscillations in the central density, during which they would continue to grow¹⁷. One black hole would probably grow much faster than the others in a runaway manner²¹. Its mass would be modest at first, perhaps $10^3 M_\odot$, and might remain considerably smaller than $3 \times 10^6 M_\odot$. This model is not ruled out, but still requires a black hole (although embedded in a dark cluster) to fit the data. A cluster of white dwarfs would evolve similarly if the white dwarfs collapse to neutron stars upon merging; if they explode as supernovae instead the cluster would either lose mass or become observable.

A cluster of brown dwarfs or planets would have a long collapse time, but a short collision time¹⁸. The outcome of a collision depends on the ratio of the collision and binding energies²²; brown dwarfs of mass $m \geq 0.05 M_\odot$ would merge with little mass loss, leading again to the growth of a massive object; planets or

small brown dwarfs would disintegrate, leading to a massive gas cloud that would be unlikely to remain dark. The constraint that the collision time be longer than the age of the galaxy rules out clusters with $m \leq 0.03 M_\odot$. It might allow¹⁸ a cluster of brown dwarfs just below the hydrogen-burning limit ($m \approx 0.08 M_\odot$) if $r_h \approx 0.08 \text{ arcsec}$, but another argument¹⁹ rules that out. The luminous stars of mass m_* that pass through the cluster would get trapped by dynamical friction in a time that depends only on the cluster density and not on m (if $m \ll m_*$), $5.5 \times 10^8 \text{ yr}$ ($M_\odot/m_*$) ($r_h/0.08 \text{ arcsec}$)^{3/2} (the collapse time given above reduced by m/m_*). The stars would sink to the centre, probably to merge into a massive object because binaries would be unable to eject stars from the deep potential well. A cluster with $r_h \approx 0.08 \text{ arcsec}$ would trap nearly 0.1% of the luminous stars, too many to be acceptable. This same argument limits the size of a bizarre dark cluster of elementary particles or tiny black holes to $r_h \leq 0.01 \text{ arcsec}$.

The simplest and most plausible interpretation of the data is thus a single massive black hole. The X-ray emission from the centre of M32 could be from accretion onto the black hole, although its flux is low enough to be from a low-mass X-ray binary²³. The absence of a larger flux is explicable³: stars will be tidally disrupted¹⁸ only once every $10^3\text{--}10^4 \text{ yr}$, and the luminosity from continuous gas accretion can be suppressed²⁴.

The arguments against a dark cluster in M32 rely mainly on the high density it would have ($\geq 10^8 M_\odot \text{ pc}^{-3}$). These arguments are weaker for the active galaxies M87 and NGC4261, for which the dark matter density required by HST gas kinematics^{25,26} is only $10^5 M_\odot \text{ pc}^{-3}$ (similar to the density of luminous material observed in some galactic nuclei and globular clusters). Stronger arguments can be made for three other black-hole candidate galaxies requiring high densities: $10^{7.3} M_\odot \text{ pc}^{-3}$ for the quiescent galaxy NGC3115, as derived from HST stellar kinematics²⁷; $10^{9.8} M_\odot \text{ pc}^{-3}$ for our own Galaxy, as derived from near-infrared velocity measurements of individual stars²⁸; and $(10^{9.6}\text{--}10^{12.6}) M_\odot \text{ pc}^{-3}$ for the active galaxy NGC4258, as derived from water maser observations^{29,30}. Together with our M32 results this provides compelling evidence for massive black holes in both active and quiescent galaxies. $\square$

Received 23 July 1996; accepted 10 January 1997.

1. Rees, M. J. *Science* **247**, 817–823 (1990).
2. Haehnelt, M. & Rees, M. J. *Mon. Not. R. Astron. Soc.* **263**, 168–178 (1993).
3. Kormendy, J. & Richstone, D. A. *Rev. Astron. Astrophys.* **33**, 581–624 (1995).
4. Tonry, J. L. *Astrophys. J.* **322**, 632–642 (1987).
5. Dressler, A. & Richstone, D. O. *Astrophys. J.* **324**, 701–713 (1988).
6. Richstone, D. O., Bower, G. & Dressler, A. *Astrophys. J.* **353**, 118–122 (1990).
7. Qian, E., de Zeeuw, P. T., van der Marel, R. P. & Hunter, C. *Mon. Not. R. Astron. Soc.* **274**, 602–622 (1995).
8. Dehnen, W. *Mon. Not. R. Astron. Soc.* **274**, 919–932 (1995).
9. Bender, R., Kormendy, J. & Dehnen, W. *Astrophys. J.* **464**, L123–L126 (1996).
10. van der Marel, R. P. et al. *Mon. Not. R. Astron. Soc.* **268**, 521–543 (1994).
11. Schwarzschild, M. *Astrophys. J.* **232**, 236–247 (1979).
12. Richstone, D. O. & Tremaine, S. *Astrophys. J.* **296**, 370–378 (1985).
13. Lauer, T. R. et al. *Astron. J.* **104**, 552–562 (1992).
14. Lugger, P. M. et al. *Astron. J.* **104**, 83–91 (1992).
15. Crane, P. et al. *Astron. J.* **106**, 1371–1393 (1993).
16. Morris, M. *Astrophys. J.* **408**, 496–506 (1993).
17. Lee, H. M. *Mon. Not. R. Astron. Soc.* **272**, 605–617 (1995).
18. Goodman, J. & Lee, H. M. *Astrophys. J.* **337**, 84–90 (1989).
19. Quinlan, G. D. *New Astron. J.* **1**, 255–270 (1996).
20. Quinlan, G. D. & Shapiro, S. L. *Astrophys. J.* **343**, 725–749 (1989).
21. Lee, H. M. *Astrophys. J.* **418**, 147–162 (1993).
22. Benz, W. & Hills, J. G. *Astrophys. J.* **323**, 614–628 (1987).
23. Eskridge, P. B., White, R. E. & Davis, D. S. *Astrophys. J.* **463**, L59–L62 (1996).
24. Narayan, R., Yi, I. & Mahadevan, R. *Nature* **374**, 623–625 (1995).
25. Harms, R. J. et al. *Astrophys. J.* **435**, L35–L38 (1994).
26. Ferrarese, L., Ford, H. C. & Jaffe, W. *Astrophys. J.* **470**, 444–459 (1996).
27. Kormendy, J. et al. *Astrophys. J.* **459**, L57–L60 (1996).
28. Eckart, A. & Genzel, R. *Nature* **383**, 415–417 (1996).
29. Miyoshi, M. et al. *Nature* **373**, 127–129 (1995).
30. Maoz, E. *Astrophys. J.* **447**, L91–L94 (1995).

Acknowledgements. We thank N. Cretton, L. Hernquist, S. Sigurdsson and S. White for collaboration in various stages of this project. The results are based on observations with the NASA/ESA Hubble Space Telescope obtained at the Space Telescope Science Institute. STScI provided support through a GO grant and through a Hubble Fellowship (R.P.v.d.M.). STScI is operated under NASA contract by the Association of Universities for Research in Astronomy, Inc.

Correspondence should be addressed to R.P.v.d.M. (e-mail: marel@sns.ias.edu).

Asymmetric synthesis by enantiomer-selective activation of racemic catalysts

Koichi Mikami & Satoru Matsukawa

Department of Chemical Technology, Faculty of Engineering,
Tokyo Institute of Technology, Meguro-ku, Tokyo 152, Japan

Asymmetric catalysis of organic reactions to provide enantio-merically enriched products is of central importance to modern synthetic and pharmaceutical chemistry^{1–3}. While non-racemic catalysts can generate non-racemic products, racemic catalysts inherently produce only a racemic mixture of chiral products. But a strategy whereby a racemic catalyst is selectively deactivated by a chiral molecule has been shown to yield non-racemic products^{4–6}. Here we describe an alternative, conceptually opposite strategy to asymmetric catalysis in which a chiral activator selectively activates one enantiomer of a racemic chiral catalyst. Our catalyst is a titanium(IV) complex for which a chiral additive acts as the chiral activator. The advantage of this approach over the deactivation strategy is that the activated catalyst can produce a greater enantiomeric excess in the products than can the enantiomerically pure catalyst on its own, as our results demonstrate.

Our strategy, and the deactivating approach, are shown schematically in Fig. 1. Our target is the catalytic asymmetric carbonylene reaction, which has emerged in recent years as one of the most efficient asymmetric carbon–carbon bond-construction processes⁷. It involves nucleophilic attack by an olefin on a carbonyl group, is catalysed by chiral Lewis acids such as binaphthol-bound titanium(IV) (BINOLate-Ti(IV)) complexes, and provides carbonylene products with high enantioselectivity⁸. Here we report the asymmetric activation of a racemic BINOLate-Ti(OPr)₂ catalyst (1) (refs 9–12 and K. B. Sharpless, personal communication) in the enantioselective catalysis of this reaction (Prⁱ denotes isopropyl).

Table 1 Carbonylene reactions with racemic BINOLate-Ti complex

Run	Activator	Yield (%)	e.e. (%)
1	None	5.9	0
2		20	0
3		38	80.8
4		52	89.8
5*		35	80.0

Conditions as in the text, unless otherwise noted.
* The reaction was carried out with 2.5 mol% of (R²)-BINOL.

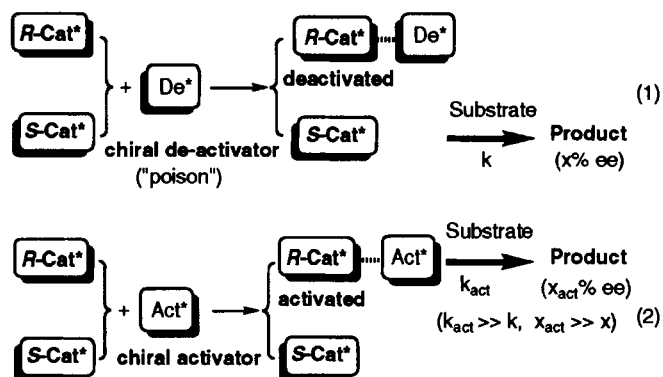


Figure 1 Concept of asymmetric activation.

The catalyst was prepared by mixing BINOL and Ti(OPr)₄ in a ratio of 1:1 in toluene for 20 min (ref. 9), and then adding 0.25–1.0 equivalent of additives such as 5,5'-dichloro-4,4',6,6'-tetramethylbiphenol (5-Cl-BIPOL) and BINOL. The carbonylene reaction was carried out *in situ* through the further addition of an olefin and glyoxylate at 0 °C to the toluene solution. The reaction was monitored by thin-layer chromatography analysis. Standard workup followed by flash chromatography afforded the carbonylene product. Chemical and optical yields of the carbonylene product were determined by gas chromatography (OV-1701, standard: benzophenone) and chiral high-performance liquid chromatography (Daicel chiral AS column) analyses, respectively.

The critical importance of the chirality of an activator was first examined by adding various activators to the racemic catalyst (1) (Table 1). Only chiral activators afforded a non-racemic product, supporting the importance of activator chirality in the selective activation of one enantiomer of the racemic chiral catalyst. Significantly, high enantioselectivities (80.8% and 89.8% enantiomeric excess, respectively, *R*) were established using a half-mole equivalent (5 mol% with respect to the substrates) of acidic (R²)-5-Cl-BIPOL and (R²)-BINOL activators even when starting from the racemic (±)-BINOLate-Ti complex (1) (10 mol%). The asymmetric activation is highlighted in a catalytic version (Table 1, run 5). A high enantioselectivity (80.0% enantiomeric excess, e.e.) is also established by adding just a catalytic amount (0.25 mole equivalent) of additional (R²)-BINOL.

Quantitative analysis of the ¹³C NMR (inverse gated ¹H-decoupling mode) spectra of the (R)-catalyst/(R)-BINOL complex suggests a molecular assembly into the hexa-coordinated and, hence, monomeric (R)-BINOLate-Ti(OPr)₂/(R)-BINOL complex (1'; Fig. 2) to show the highly symmetrical nature (δ_c 153.5 (2C), 162.6 (2C) and 82.8 (2C)); free BINOL shows a hydroxy-carbon signal at 153.1 p.p.m. (s), while BINOLate-Ti(OPr)₂ (1) shows the BINOLate hydroxy-carbon and isopropoxy signals at 159.8 (s) and 80.9 (s) p.p.m., respectively. A similar hexa-coordinated and monomeric (R)-BINOLate-Ti(OPr)₂/(R)-5-Cl-BIPOL complex (1'; Fig. 2) can also be identified by ¹H NMR analysis (6.54 (2H, s) p.p.m.); free 5-Cl-BIPOL shows the H³ signal at 6.51 p.p.m. (s), while 5-Cl-BIPOLate-Ti(OPr)₂ shows the H³ signal at 6.78 p.p.m. (s).

The activation of the (R¹)-BINOLate-Ti(IV) catalyst (1) was then confirmed by the addition of 1 mole equivalent of (R²)-BINOL (Table 2). The reaction proceeded smoothly to provide the carbonylene product in an increased chemical yield (66.0% and 82.1%) along with an enhanced level of enantioselectivity (97.2% e.e. and 96.8% e.e.) (Table 2, runs 3 and 4) compared to the values obtained using the (R¹)-BINOLate-Ti(IV) catalyst (1) without additional acidic diols (19.8%, 94.5% e.e.) (Table 1, run 1). Thus, the (R¹)-BINOLate-Ti(IV) catalyst (1) is activated by the addition of

another 1 mole equivalent of (R^2)-biphenol, as is evidenced by the increased chemical and optical yields. Comparing the results of enantiomer-selective activation (Table 1, run 4 versus Table 2, runs 1 and 4), the reaction catalysed by the (R^1)-catalyst/(R^2)-BINOL complex ($1'$) is calculated to be 26 times as fast as that catalysed by the original catalyst (1), (S^1)- 1 in the racemic case. Kinetic studies confirm this, showing that the reaction catalysed by the (R^1)-catalyst/(R^2)-BINOL complex ($1'$) is 26 times as fast as that catalysed by the (R^1)-catalyst (1) (Fig. 3, equation (3)) according to a rapid-quench gas chromatography analysis (Table 2, run 2 versus run 5). Sharpless *et al.* have emphasized the importance of 'chiral ligand acceleration'¹³, however, in the construction of a chiral catalyst from an achiral pre-catalyst. These results imply that the racemic ($\pm$)-BINOLate-Ti(IV) catalyst (1) and a half-mole equivalent of (R^2)-BINOL are assembled preferentially into the (R^1)-BINOLate-Ti(IV)/(R^2)-BINOL complex ($1'$) in addition to the remaining (S^1)-BINOLate-Ti(IV) catalyst (1) (equation 3).

In contrast, the enantiomeric form of the additional chiral ligand ((S^2)-BINOL) activates the (R^1)-BINOLate-Ti(IV) (1) catalyst less than (R^2)-BINOL, thus providing the carbonylene product in lower chemical (48.0%) and optical (86.0% e.e., R) yields (Table 2, run 7). The racemic combination of (R)- and (S)-BINOL affords the enantio-enriched product by the step-wise addition of the enantiomers. Therefore, the formation of the spiro titanium complex [BINOLate]₂-Ti (5) or the ligand exchange are highly unlikely. In fact, the formation of spiro complex (5) was not observed even after azeotropic removal of isopropanol from 2:1 mixture of BINOL and Ti(OPr)₄ in toluene.

We then explored the effect of using racemic BINOL as an activator. Racemic BINOL was added to the (R^1)-BINOLate-Ti(IV) (1) catalyst to give an enhanced level of enantioselectivity (95.7% e.e., 69.2%) (Table 2, run 9) relative to that obtained by the original catalyst (1) (94.5% e.e., 19.8%) (run 1). The (R^1)-catalyst/(R^2)-BINOL complex was thus activated by a factor of 8.8 relative to the

Table 2 Carbonylene reactions with enantio-pure (R)-BINOLate-Ti complex

Run	Chiral activator	Time (min)	Yield (%)	e.e. (%)
1	None	60	19.8	94.5
2		1	1.8	94.5
3	(R^2)-5-Cl-BIPOL	60	66.0	97.2
4	(R^2)-BINOL	60	82.1	96.8
5		1	41.1	96.8
6		0.5	24.0	96.9
7	(S^2)-BINOL	60	48.0	86.0
8		0.5	2.6	86.0
9	($\pm^2$)-BINOL	60	69.2	95.7

Conditions as in the text.

(R^1)-catalyst/(S^2)-BINOL complex. The reaction catalysed by the (R^1)-catalyst/(R^2)-BINOL complex is in fact 9.2 times as fast as that catalysed by the (R^1)-catalyst/(S^2)-BINOL complex (Fig. 3, equation (4)) according to a rapid-quench gas-chromatography analysis (runs 6 versus run 8). Thus, an enhanced level of enantioselectivity (95.7% e.e.) is obtained with racemic BINOL as the chiral activator.

The great advantage of asymmetric activation is exemplified in the Diels–Alder reaction with glyoxylate (Fig. 4), where the enantio-pure BINOLate-Ti(IV) catalyst (1) provides only low (5% e.e.) enantioselectivity. By adding 0.5 mole equivalent of chiral activator, however, the racemic catalyst 1 affords an enantiomerically amplified (50% e.e.) product. The high enantioselectivity obtained using

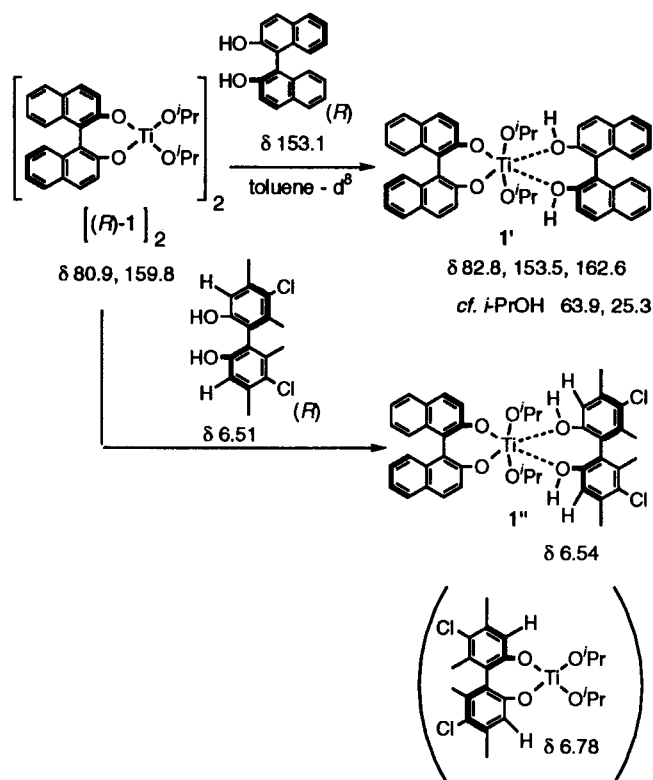


Figure 2 ^{13}C NMR (inverse gated ^1H -decoupling mode) and ^1H NMR data for (R)-BINOL-Ti(IV)/(R)-BINOL and (R)-BINOL-Ti(IV)/(R)-5-Cl-BIPOL complexes, respectively.

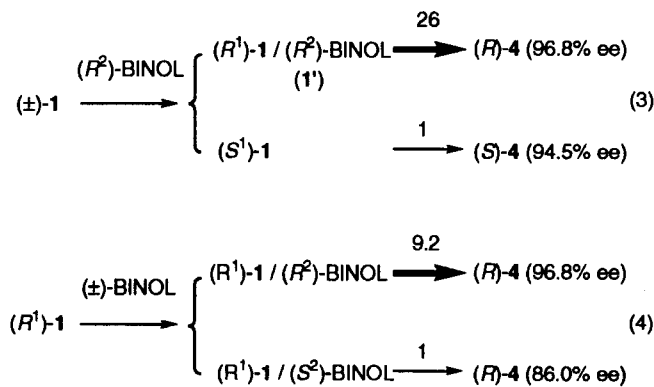


Figure 3 Kinetic studies in carbonylene reactions with racemic and single-enantiomer BINOLate-Ti complexes.

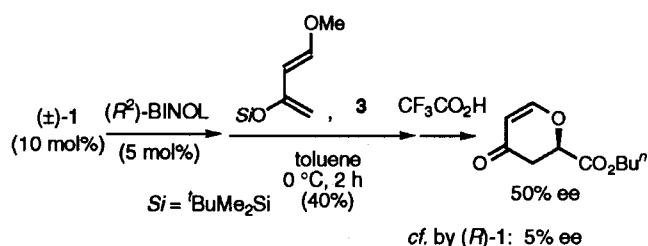


Figure 4 The Diels–Alder reaction with racemic BINOLate-Ti complex.

the racemic catalyst (1) can be regarded as a kind of positive nonlinear effect^{14,15} (asymmetric amplification)^{16–18}.

We believe that our asymmetric activation strategy will provide a generally effective strategy for racemic catalysts used in catalytic asymmetric processes^{19,20}. □

Received 11 September; accepted 10 December 1996.

1. Doyle, M. (ed.) *Advances in Catalytic Processes: Asymmetric Chemical Transformations* Vol. 1 (JAI, Greenwich, CT, 1995).
2. Noyori, R. *Asymmetric Catalysis in Organic Synthesis* (Wiley, New York, 1994).
3. Kagan, H. B. *Comprehensive Organic Chemistry* Vol. 8 (Pergamon, Oxford, 1992).
4. Alcock, N. W., Brown, J. M. & Maddox, P. J. *J. Chem. Soc. Chem. Commun.* 1532–1533 (1986).
5. Maruoka, K. & Yamamoto, H. *J. Am. Chem. Soc.* **111**, 789–790 (1989).
6. Faller, J. W. & Parr, J. *J. Am. Chem. Soc.* **115**, 804–805 (1993).
7. Mikami, K. & Shimizu, M. *Chem. Rev.* **92**, 1021–1050 (1992).
8. Mikami, K., Terada, M. & Nakai, T. *J. Am. Chem. Soc.* **112**, 3949–3954 (1990).
9. Martin, C. A. thesis, MIT (1988).
10. Wang, J. T., Fan, X., Feng, X. & Qian, Y. M. *Synthesis* 291–292 (1989).
11. Keck, G. E., Tarbet, K. H. & Geraci, L. S. *J. Am. Chem. Soc.* **115**, 8467–8468 (1993).
12. Faller, J. W., Sams, D. W. & Liu, X. *J. Am. Chem. Soc.* **118**, 1217–1218 (1996).
13. Berrisford, D. J., Bolm, C. & Sharpless, K. B. *Angew. Chem. Int. Edn Engl.* **34**, 1059–1070 (1995).
14. Guillauneux, D., Zhao, S.-H., Samuel, O., Rainford, D. & Kagan, H. B. *J. Am. Chem. Soc.* **116**, 9430–9439 (1994).
15. Mikami, K. & Terada, M. *Tetrahedron* **48**, 5671–5680 (1992).
16. Kagan, H. B. et al. *J. Am. Chem. Soc.* **108**, 2353–2357 (1986).
17. Noyori, R. & Oguni, N. et al. *J. Organomet. Chem.* **382**, 19–37 (1990).
18. Noyori, R. & Kitamura, M. *Angew. Chem. Int. Edn Engl.* **30**, 49–69 (1991).
19. Mikami, K. & Matsukawa, S. *Tetrahedron Asymmetry* **6**, 2571–2574 (1995).
20. Mikami, K. & Matsukawa, S. *Enantiomer* **1**, 69–73 (1996).

Acknowledgements. We thank H. B. Kagan and J. M. Brown for stimulating and useful discussions.

Correspondence and requests for materials should be addressed to K.M.

Acoustic observations of heat content across the Mediterranean Sea

U. Send*, G. Krahmann*, D. Mauuary*, Y. Desaubies†, F. Gaillard†, T. Terre†, J. Papadakis‡, M. Taroudakis‡, E. Skarsoulis‡ & C. Millot§

* Institut für Meereskunde, Düsterbrookweg 20, 24105 Kiel, Germany

† Laboratoire de Physique des Océans, IFREMER, Plouzané, France

‡ Institute for Applied and Computational Mathematics, Foundation for Research and Technology Hellas, PO Box 1527, 71110 Heraklion, Crete, Greece

§ Centre d'Océanologie de Marseille, Antenne COM, CNRS, BP 330-83507, La Seyne-sur-Mer, France

The ability to monitor the heat content of oceans over long distances is becoming increasingly important for understanding the role of oceans in climate change, for determining the variability of the state of the oceans, for operational ocean observing systems, and for studying large-scale ocean processes such as water-mass formation. Although the properties of the upper layers of the ocean can be routinely measured on large scales by satellite remote sensing (providing altimetric and infrared data) and with expendable probes dropped from commercial vessels, the deep interior of the ocean is more difficult to monitor. Ocean acoustic tomography¹ is a promising technique for such applications, as it has the potential to provide systematic, instantaneous and repeated measurements of the ocean interior over large parts of an ocean basin. Here we demonstrate the capability of this technique for measuring the heat content across an entire (albeit small) ocean basin—the western Mediterranean Sea.

Minimal absorption in the ocean allows low-frequency acoustic transmissions over long horizontal distances. Measuring the travel time of the sound then yields large-scale integrals of the (inverse) sound speed, which are tightly related to heat content. The most ambitious application of this approach is the current ATOC project

(Acoustic Thermometry of Ocean Climate)¹⁵ which aims to observe climate signals in the deep ocean by basin-scale acoustic transmission in the Pacific Ocean. A prior demonstration of heat content measurements using tomography on a 1,000 km scale was given in ref. 3. There, data collected over a four-month period along a triangle in the open Pacific yielded good measurements of the heat content, but a heat budget combining heat content and surface heat fluxes could not be balanced, because of horizontal advection effects in that open ocean environment.

Here we present preliminary results from a nine-month-long tomography experiment involving seven moorings which was conducted in the western Mediterranean Sea. A combination of aspects make this application unique and particularly relevant to the issues mentioned above. The western Mediterranean is a nearly closed basin, and a large part of it was covered with the distributed network of seven instruments. This allowed estimation of the various terms of the heat budget during a seasonal cycle. A direct comparison with the surface heat fluxes over the basin, which have been the topic of a variety of recent studies^{4–6}, is therefore possible. The Mediterranean also is a convenient test 'laboratory', where most of the oceanographic processes of the major oceans can be found, while it is more accessible, controlled (nearly closed), and of smaller size. For example, a climate-like deep warming trend has been observed in some part of the basin, but it is not clear how widespread this signal is and whether it is caused by changing atmospheric conditions^{7,8}. Another advantage is that the scales of the basin are such that it can be covered with a network of standard tomographic instruments. In addition, various numerical models of the region exist already for assimilation studies.

The experiment, called THETIS-2, was conducted in order to demonstrate the capabilities of tomography for basin-scale measurements and heat flux calibrations, to provide a data set for assimilation studies, and to prove the concept for long-term acoustic monitoring in climate studies or operational systems. Figure 1 shows the sampling of the western Mediterranean by the mooring array of tomography instruments that was deployed for the THETIS-2 experiment from January until October 1994 (covering the extrema of the annual heat content cycle). The measurement principle is the standard acoustic tomography approach¹, where the discrete acoustic arrivals result from different depth penetration of the corresponding ray paths and thus provide a vertical sampling of the ocean (Fig. 1 bottom). For initialization and verification, a substantial number of hydrographic profiles (temperature and salinity) were collected in the basin covered by the experiment, and an XBT (expendable temperature probe) line was occupied every two weeks using a commercial ship. The transmission path from H to W3 (Fig. 1 top) was specifically aligned exactly with this XBT track, in order to allow detailed comparisons.

After standard tomographic signal processing², Doppler phase drift compensation, and clock and mooring motion correction, the acoustic travel times represent a high-quality measure for the temperature integrated along the ray paths across the basin. We present preliminary analyses of total heat content along three of the sections and of a three-dimensional average of the basin heat content. At present, only single deep-turning ray paths were used in these analyses (Fig. 1), as such paths sample most of the water column. Numerical simulations with historical seasonal vertical profiles, vertical empirical orthogonal functions and typical oceanographic perturbations for each tomography section were used to establish a (nonlinear) relation between the travel times and the total heat content of the section including error estimates. Figure 2 gives the results of this relation yielding the heat content of the section H–W3 from the travel times of a deep-turning ray, compared to estimates of the same quantity from the CTD and XBT surveys. The figure shows a high degree of agreement between the two types of data, and the deviations are in the range of the calculated error bars.

The main signal seen is, as expected, the seasonal cycle of heating and cooling, but there are significant deviations from the historical mean annual cycle, which is also plotted in Fig. 2. The departures of the acoustic measurements for 1994 from the mean cycle are about 10% in terms of overall amplitude for this particular section—on other sections the difference is larger or has the opposite sign. In addition to the annual cycle there are distinct features and events of limited duration. These are also more pronounced on other sections analysed (see Fig. 3) and some of them are undoubtedly related to the significant mesoscale variability revealed by the XBT sections. We have processed two more sections the same way (H–W1 and H–W5; Fig. 1), and combined them to give an optimal linear estimate of the horizontally averaged (three-dimensional) heat content of much of the basin (1–9°E), using horizontal covariance scales computed from historical data and from a numerical model. The

resulting curve for the 1–9°E basin-mean heat content is shown in Fig. 3 top, along with the three individual time series that were used to construct it. Rather large regional differences are visible with pronounced local events, as well as their signature in the optimally weighted average. The three-dimensional average is only preliminary, as the acoustic estimate will improve as more sections are included. However, this is the first time-series of instantaneous three-dimensional heat content estimates for the Mediterranean, and a comparison with surface heat fluxes is already possible.

Surface heat fluxes represent a major forcing for the ocean, yet they still have insufficient accuracy for many oceanographic applications. For example, observations of the heat transport through the Strait of Gibraltar dictate that the average surface heat loss in the Mediterranean should be about -5 W m^{-2} (ref. 9), but direct calculations for the heat flux from ship measurements have ranged from 0 to $+29 \text{ W m}^{-2}$ (refs 4–6). Gilman and Garrett⁶ suggested various corrections to achieve agreement with the Gibraltar measurements in the mean, but calibrations or corrections for the time-varying parts are even harder to perform. With our heat storage observations this can be attempted, because for a nearly closed basin, the time variations in the total heat content must equal the time integral of the surface heat fluxes plus that of the heat transport through the openings (straits).

We have obtained ECMWF (European Center for Medium Range Forecasting) surface heat fluxes for the Mediterranean for the entire year 1994, and integrated them in time and over the western basin between the Straits of Gibraltar and Sicily, as well as the above 1–9°E subregion. After correction for heat transport through the straits, the result can be compared to the observed heat storage time series (Fig. 3 bottom). We see that in spite of the approximations used, the amplitude of the seasonal cycle agrees to within ~3%,

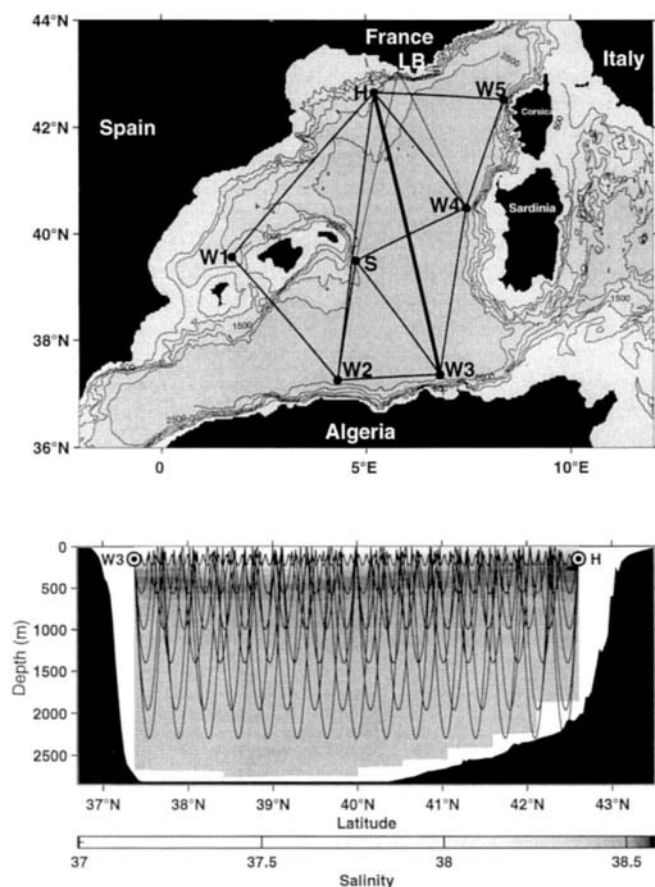


Figure 1 Horizontal and vertical sampling of the western Mediterranean in the THETIS-2 experiment. Top panel, network of tomography sections in the basin. A 250-Hz HLF-5 sound source at location H sent sound waves into the entire basin under investigation. The remaining instruments (W1–W5 and S) were 400-Hz transceivers (combined transmitters and receivers). All systems except S had dual-channel receivers to record both the 250 Hz and 400 Hz signals. The seven instruments were operational for the time period deployed, and each transmission path shown had a large number of well resolved stable acoustic arrivals (corresponding to the ray paths in the bottom panel) in at least one direction. The heavy line along H–W3 marks the XBT section occupied every 2 weeks; and LB denotes a land-based receiving/monitoring hydrophone. Bottom panel, vertical distribution of some discrete ray paths which the sound energy could take from mooring H to W3, superimposed here on the salinity field in practical salinity units (shading) to indicate the typical scales of oceanic variability. (Data shown here are for October 1994) Each ray path samples a different depth range along the section, and can be identified in the data because it has a unique discrete travel time. This allows a vertical sampling of the section.

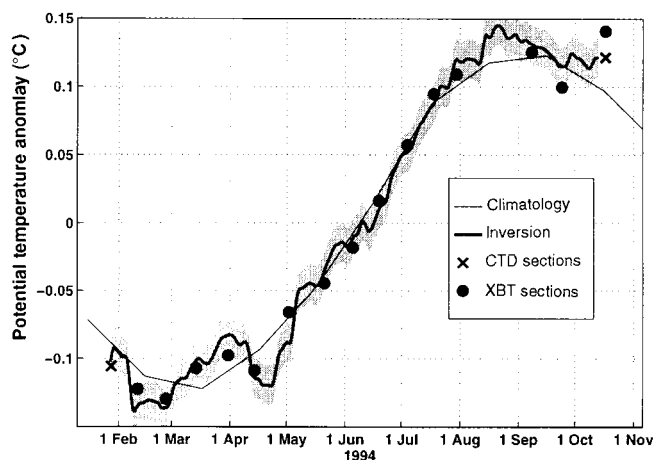


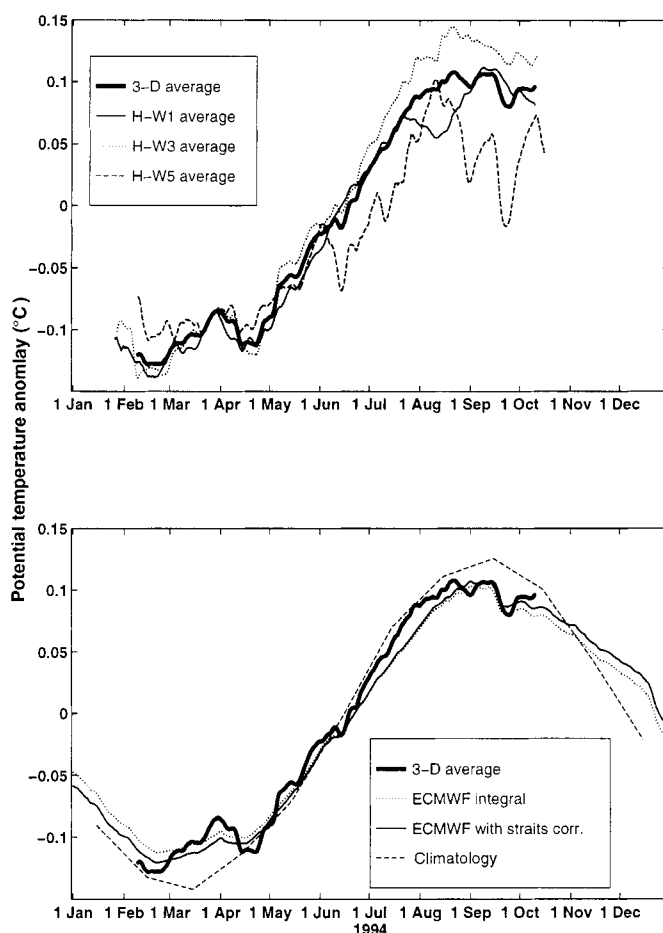
Figure 2 The heavy black line shows the acoustic average of potential temperature (relative to 13.111 °C) over the 0–2,000 m layer and over the 600-km-long section from H to W3 (see Fig. 1), using the travel-time measurements for a deep-turning ray path. Dots and crosses are estimates of the same quantity from CTD (lowered conductivity–temperature–depth sensors) and XBT surveys along the same line. The error interval (shaded) indicates the uncertainty of the acoustic estimates based on scatter in the travel time/heat content relation and on uncertainty in the travel-time measurements. The error of the CTD and XBT averages is smaller than this, but we note that the XBTs only sample to 800 m depth. Also shown is the annual mean curve for the 0–2,000 m average temperature along the section from the climatology (thin black). This mean seasonal cycle was constructed from a comprehensive climatology, based on 85,000 MBT/XBT (mechanical and expendable bathythermographs)/CTD and bottle sample profiles. Averaged over the whole basin, the heat content amplitude of the climatology is in very good agreement with mean surface heat fluxes calculated from the COADS (comprehensive ocean atmosphere data set)¹² data for the years 1946–1992 after Gilman and Garrett².

Figure 3 Top panel, individual estimates of 0–2,000 m mean temperature anomaly along the acoustic sections H–W1, H–W3 and H–W5, processed like the data in Fig. 2 (thin lines). The optimal linear estimate from these for the three-dimensional average over the 1–9° E part of the basin is shown by the heavy curve (error $\sim 0.015^\circ\text{C}$). Bottom panel, the three-dimensional average heat content for 1994 from the top panel (heavy line) together with the climatological curve (dashed line) and the time-integral of the ECMWF surface heat fluxes for the 1–9° E sub-basin (dotted line). Strictly, the surface fluxes cannot be integrated over a sub-region of the basin, as the ocean may transport the heat out of this area. We tested the relation by comparing the 1–9° E sub-region with the full basin in climatological heat-content data and in ECMWF heat flux integrals. The entire basin had an amplitude larger by 6% and 9% respectively. Therefore we have adjusted both the full-basin ECMWF integral and climatological seasonal amplitudes by a mean factor of 7.5% to represent data for the 1–9° E sub-region. At a later stage, we expect to perform this 'extrapolation' dynamically via assimilation into numerical models. For comparison with the heat-content observations, the heat transport through the openings of the basin should also be added to the surface heat flux. Using the historical data and results from refs 9, 13 and 14, the estimates for the Strait of Gibraltar and Strait of Sicily are $3.3 \pm 0.8^\circ\text{C Sv}$ and $2.1 \pm 1.3^\circ\text{C Sv}$, respectively ($1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$), yielding a net heat transport into the western Mediterranean of $1.2 \pm 1.5^\circ\text{C Sv}$ which is equivalent to an area-mean surface heat flux of $5.7 \pm 7 \text{ W m}^{-2}$. Between the extrema of the seasonal cycle (March–September) this gives an extra heating of $0.011 \pm 0.013^\circ\text{C}$ averaged over the upper 2,000 m of the water column. This extra heating should be added to the temperature changes predicted from the ECMWF heat flux integral. The thus corrected curve is shown as a thin solid line.

which corresponds to a mean heat flux of $\sim 4 \text{ W m}^{-2}$. Some of this agreement may be fortuitous, as the error in the tomography estimates is $\sim 0.015^\circ\text{C}$ in this preliminary analysis. Combined with the uncertainty in the strait transports, the actual error bar for the three-dimensional average is about four times larger than the current mismatch. In addition to the mean seasonal heating amplitude, there is also some agreement visible for single events, especially the strong cooling in late September. Note that the climatology shows a seasonal amplitude $\sim 20\%$ higher than the ECMWF flux integral. More detailed estimates, after complete processing of the tomography data available and with improved strait transport measurements currently underway, could provide a severe test for surface heat flux formulae.

Another unique aspect of this data set is the fact that it provides a good unaliased spatial integral coverage and long time series for an ocean basin for which a number of realistic numerical models with current assimilation studies exist. Numerical models can benefit from assimilating real data about the state of the ocean, either to determine and correct model deficiencies or to initialize operational forecasting simulations, and assimilations will soon be routinely performed with, for example, satellite altimetry data. The expectation is that large-scale internal temperature observations as obtained here will complement the satellite surface data and will efficiently constrain models and/or forcing fields¹⁰. Even if the absolute heat content of a model is not dynamically relevant, its time evolution and the large regional differences seen in Fig. 3 top are essential observables that a model should reproduce—not necessarily at single points but in an area-mean. The first joint assimilation of real tomography and altimetry data has been performed using the present data in a simple application and is presented in an accompanying paper¹¹. In that study, a time-evolving circulation state for the basin was found that is consistent with the model dynamics, the altimetry and the tomography integrals.

The results presented here have only analysed a subset of the available data. Measurements are available along all the sections shown in Fig. 1, and taken altogether the accuracy of the area-mean or the spatial resolution will significantly increase. Also, here we have not exploited the depth resolution contained in the sampling



by ray paths penetrating to various depths (Fig. 1). This work is in progress, and has been successfully performed already for the H–W3 section. It is clear that separate heat-content timeseries can be obtained for the depth ranges corresponding to the different water masses (surface layer, Levantine Intermediate Water, and deep water).

Most of the moorings in THETIS-2 were located just off steep topography near a coast, where the same measurements could be performed with land-cabled instruments. Thus the THETIS-2 project has also demonstrated that the large-scale fluctuations and evolution of a basin like the Mediterranean could be sampled and monitored with existing means by acoustic transmissions between shore-based stations. □

Received 25 June 1996; accepted 2 January 1997.

1. Munk, W. & Wunsch, C. *Deep-Sea Res.* **26**, 123–161 (1979).
2. Munk, W., Worcester, P. F. & Wunsch, C. *Ocean Acoustic Tomography* (Cambridge Univ. Press, 1995).
3. Dushaw, B. D., Worcester, P. F., Cornuelle, B. D. & Howe, B. M. *J. Phys. Oceanogr.* **23**, 2650–2666 (1993).
4. Bunker, A. F., Charnock, H. & Goldsmith, R. A. *J. Mar. Res. Suppl.* **40**, 73–84 (1982).
5. Garrett, C., Outerbridge, R. & Thompson, K. *J. Clim.* **6**, 900–910 (1993).
6. Gilman, C. & Garrett, C. *J. Geophys. Res.* **99**, 5119–5134 (1994).
7. Bethoux, J. P., Gentili, B., Raunet, J. & Tailliez, D. *Nature* **347**, 660–662 (1990).
8. Rohling, E. J. & Bryden, H. L. *J. Geophys. Res.* **97**, 11191–11198 (1992).
9. Macdonald, A. M., Candela, J. & Bryden, H. L. in *Seasonal and Interannual Variability of the Western Mediterranean Sea* (ed. La Violette, P. E.) 13–32 (Am. Geophys. Union, Washington DC, 1994).
10. Munk, W. & Wunsch, C. *Phil. Trans. Royal Soc. A* **307**, 439–464 (1982).
11. Menemenlis, D., Webb, T., Wunsch, C., Send, U. & Hill, C. *Nature* **385**, 618–621 (1997).
12. Woodruff, S. D., Slutz, R. J., Jenne, R. L. & Steurer, P. M. *Bull. Am. Meteorol. Soc.* **68**, 1239–1250 (1987).
13. Manzella, M. R., Gasparini, G. P. & Astraldi, M. *Deep-Sea Res.* **35**, 1021–1035 (1988).
14. Astraldi, M., Gasparini, G. P., Sparnocchia, S., Moretti, M. & Sansone, S. *Bull. Inst. Oceanogr.* **17**, 95–115 (1996).
15. Munk, W. *J. Acoust. Soc. Am.* **100**, 2580 (1996).

Acknowledgements. The authors are members of the THETIS-2 group. This work was supported by the EU MAST-2 programme.

Correspondence should be addressed to U.S. (e-mail: usend@ifm.uni-kiel.de).

Basin-scale ocean circulation from combined altimetric, tomographic and model data

Dimitris Menemenlis^{*}, Tony Webb[†], Carl Wunsch^{*}, Uwe Send[‡] & Chris Hill^{*}

^{*} Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[†] Civil Engineering, University College, UNSW, ADEA, Northcott Drive, Campbell, ACT, 2601, Australia

[‡] Institut für Meereskunde, Regionale Ozeanographie, Duesternbrooker Weg 20, 24105 Kiel, Germany

The ocean stores and transports vast quantities of heat, fresh water, carbon and other materials, and its circulation plays an important role in determining both the Earth's climate and fundamental processes in the biosphere. Understanding the development of climate and important biological cycles therefore requires detailed knowledge of ocean circulation and its transport properties. This cannot be achieved solely through modelling, but must involve accurate observations of the spatio-temporal evolution of the global oceanic flow field. Estimates of oceanic flow are currently made on the basis of space-borne measurements of the sea surface, and monitoring of the ocean interior. Satellite altimetry and acoustic tomography are complementary for this purpose¹, as the former provides detailed horizontal coverage of

the surface, and the latter the requisite vertical sampling of the interior. High-quality acoustic-tomographic² and altimetric³ data are now available to test the combined power of these technologies for estimating oceanic flows. Here we demonstrate that, with the aid of state-of-the-art numerical models, it is possible to recover from these data a detailed spatio-temporal record of flow over basin-scale volumes of fluid. Our present results are restricted to the Mediterranean Sea, but the method described here provides a powerful tool for studying oceanic circulation worldwide.

The ocean fluctuates on a wide range of spatial and temporal scales⁴. The measured potential-energy spectrum of the circulation is mostly 'red', that is, the energy density increases with increasing spatial and temporal scales, but with a marked peak at the annual cycle. However, the kinetic-energy spectrum is dominated by mesoscale eddies with horizontal scales of 100 km or so, and periods of weeks to months. Therefore, studies of the large-scale movement of the sea need to resolve the mesoscale eddies and other narrow features of the circulation such as boundary-layer jets which are responsible for a large fraction of the heat and property transport. To resolve these features, the ocean would have to be instrumented roughly every 50 km horizontally and at some 20 levels vertically. The requirements for global coverage at this resolution are prohibitive.

Here we focus on the use of tomographic and altimetric observation systems which, in combination with more traditional ocean sampling methods, go far towards solving the global observation problem. Satellite altimetry measures the sea surface elevation relative to the geoid, that is, relative to the particular gravitational equipotential to which the sea surface would conform if it were at rest with no forces acting on it (other than gravity). Changes in

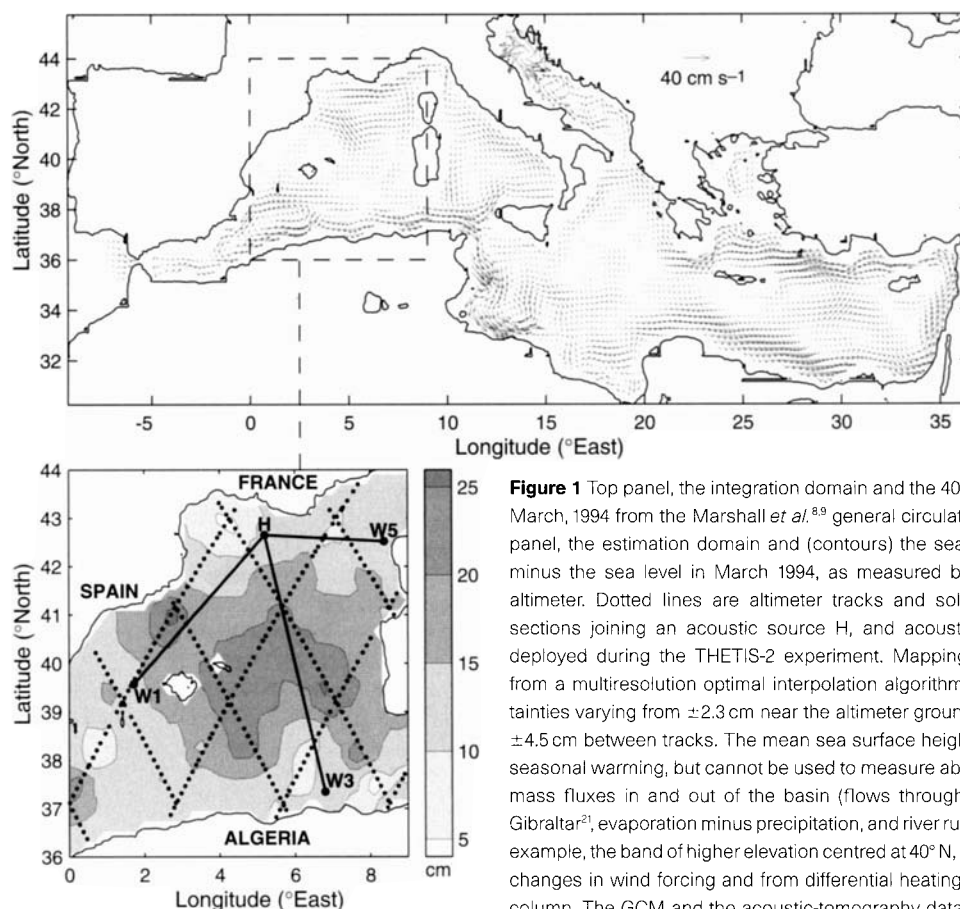


Figure 1 Top panel, the integration domain and the 40-m horizontal velocity on 1 March, 1994 from the Marshall *et al.*^{8,9} general circulation model (GCM). Bottom panel, the estimation domain and (contours) the sea level in September 1994 minus the sea level in March 1994, as measured by the TOPEX/POSEIDON altimeter. Dotted lines are altimeter tracks and solid lines are tomographic sections joining an acoustic source H, and acoustic receivers W1, W3, W5 deployed during the THETIS-2 experiment. Mapping of the altimeter data is from a multiresolution optimal interpolation algorithm²⁰. Estimates have uncertainties varying from ± 2.3 cm near the altimeter ground tracks to a maximum of ± 4.5 cm between tracks. The mean sea surface height increase is indicative of seasonal warming, but cannot be used to measure absolute heating because of mass fluxes in and out of the basin (flows through the Straits of Sicily and Gibraltar²¹, evaporation minus precipitation, and river runoff). Relative changes, for example, the band of higher elevation centred at 40° N, contain contributions from changes in wind forcing and from differential heating and cooling of the water column. The GCM and the acoustic-tomography data are used to separate the respective contributions to sea-level anomaly of wind and heat content.

elevation result from local exchange of mass and heat with the atmosphere through the sea surface, and from lateral water movements. In general terms, surface elevation and its slope provide a dynamical surface boundary condition on the general circulation.

The second observational system, ocean acoustic tomography, integrates the oceanic state along many paths through a volume of fluid by transmitting sound pulses from sources to receivers. Perturbations in travel time of acoustic pulses are dominated by temperature perturbations, and to a lesser degree by perturbations in density and velocity⁵. Both altimetric and tomographic measurements have been compared in detail (where possible) to more conventional oceanographic measurements and have been shown to be fully consistent with them^{2,6}.

We used altimetric data⁷ from the tracks shown in Fig. 1 for the period 3 October, 1992 to 25 November, 1994. The measured sea-surface height change during a six-month period, March to September 1994, concurrent with the tomographic measurements is also shown in Fig. 1. Because of the residual geoid errors and because mass is not conserved in the western Mediterranean basin, the altimeter data do not directly reflect the mean circulation or the basin-averaged changes in heat content⁶.

We also use the analysis of THETIS-2 acoustic tomography data shown in Fig. 2 and consisting of depth-averaged (0–2,000 m) and

path-averaged potential temperature along the three sections marked on Fig. 1. The tomographic data, over the instrumented period from February to October 1994, show, in addition to the expected seasonal warming through summer, a marked difference between the mid-basin section (H-W3) and the northern sections—a difference that mirrors the relative amplitude variations from the altimetry.

To recover an ocean state which is dynamically consistent with the observations, we use the general circulation model (GCM) of Marshall *et al.*^{8,9}, which is here integrated for the entire Mediterranean basin with realistic topography, 0.25° horizontal grid spacing, and 20 vertical levels (Fig. 1). Twice-daily surface wind-stress boundary conditions are from the European Center for Medium Range Weather Forecasts (ECMWF), and heat and freshwater fluxes are provided by a relaxation term at the surface to seasonal climatologies¹⁰ of temperature and salinity in order to mimic seasonal heat and freshwater fluxes. An Atlantic box is added to simulate the Gibraltar throughflow. The GCM includes a simple convective overturning scheme permitting buoyancy-driven deep-water formation to take place, but lacks a wind-driven mixed layer. In this configuration the GCM reproduces a wind-driven circulation consistent on the large scales with that obtained by previous modelling studies^{11,12}, but fails to capture the detailed thermo-

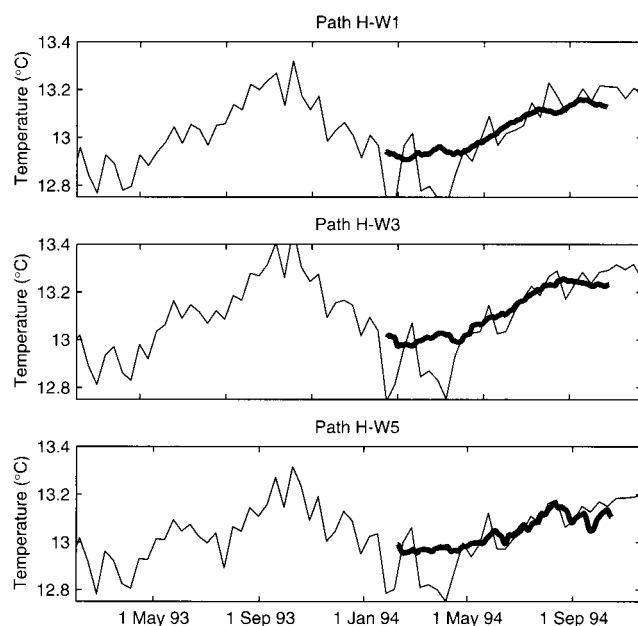


Figure 2 Depth-averaged (0–2,000 m) and path-averaged potential temperature measured using ocean acoustic tomography (thick lines) and estimated from the altimeter data (thin lines) along the three sections marked on Fig. 1. Tomographic heat content estimates are from inversions using a single deep-diving acoustic ray. The estimated uncertainty is $\pm 0.013^\circ\text{C}$. Section H-W3 was sampled on a regular basis with expendable bathythermographs during the duration of the experiment and the agreement with the inversions is to within measurement uncertainty. The data show a strong seasonal heating which is mainly confined to the top 50 m and is consistent with ECMWF heat flux analyses over the western Mediterranean². Altimeter heat content estimates are made under the assumption that the sea-level anomaly signal results from uniform heating or cooling of the top 50 m. Differences between the tomographic and the altimetric heat content estimates result from physical processes other than heating or cooling of the surface layer, for example, the response of the sea-level anomaly to changes in atmospheric pressure, wind stress, and freshwater fluxes. A key contribution of the tomographic data is the ability to estimate the heat content part of the altimetric sea-level anomaly signal.

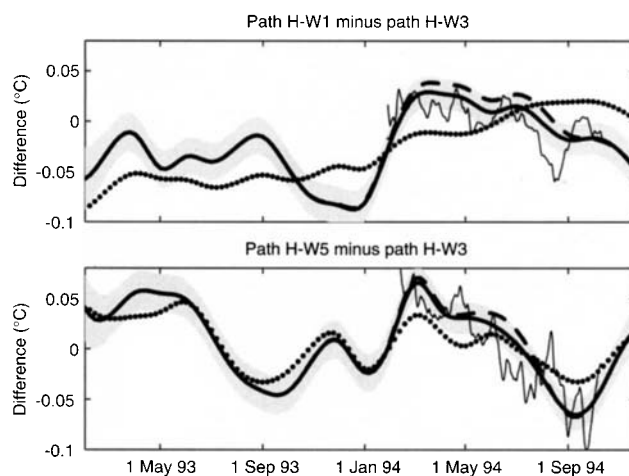


Figure 3 Difference between depth-averaged (0–2,000 m) and path-averaged potential temperature along the tomographic sections. The dotted line is the GCM estimate; the dashed line indicates the GCM/altimeter combination; the thick solid line represents the GCM/altimeter/tomography combination with the shaded area indicating the standard error of this last estimate; the thin solid line is the tomographic data; and circles represent altimeter data converted to depth-integrated temperature as in Fig. 2. GCM and data biases have been removed as discussed in the text. The GCM predictions differ significantly from the tomographic data because of the lack of realistic buoyancy forcing at the surface. However, the GCM/altimeter combination successfully recovers the time-evolving spatial anomaly of heat content, as seen through the calibration against tomographic data. The addition of the tomographic data does not change the estimates significantly, but the uncertainty is reduced by 17%, from $\pm 0.018^\circ\text{C}$ when there are no tomographic data to $\pm 0.015^\circ\text{C}$ during the period when tomographic data are available. Remaining differences between the estimates and the tomographic data are believed real—representing the still-inadequate resolution of the GCM and of the estimation method.

dynamic surface forcing and seasonal heating and cooling of the surface layer. The estimation scheme, described below, compensates in part for these and other GCM deficiencies, as well as the shortcomings of the data.

The problem of combining data and models is analogous to, but distinct in practice, from that used to depict the atmosphere for purposes of numerical weather forecasting. The spatial scales of oceanic variability are much smaller than in the atmosphere, while the temporal scales are much longer. Furthermore, we are not (yet) interested primarily in forecasting, but in understanding, so that the estimation ('assimilation') methods have the goal of the most accurate state estimate, rather than the most accurate forecast. We thus exploit the ability to use formally future data, that is, estimates at any given moment are influenced by all the observations, past, present, and future to the estimation time period.

A least-squares fit of the available data is made to a reduced-state, linear model. The reduced state describes large-scale, low-frequency temperature (density) differences between the GCM and the observations, and is represented by four vertical modes, 1° -sampling in the horizontal, and 1-month time steps. This state reduction is required to make the problem tractable with the available computational resources. The reduced-state linear model is determined by systematically perturbing the GCM and using the resulting anomalies to deduce a state transition matrix for the physical processes governing the perturbations¹³. It is assumed, and tested *a posteriori*, that the GCM is sufficiently close to the observations for the linearization to be adequate. The least-squares problem is solved using a sequential algorithm—the Kalman filter followed by the Rauch–Tung–Striebel smoother¹⁴—to obtain estimates of the reduced state and the associated uncertainty covariance matrix. Discrepancies between the GCM and observations of freshwater fluxes, external modes, and other physical processes not resolved by the reduced state are part of the error budget in this analysis.

Two difficulties are encountered. First, it is necessary to specify

second-order error statistics for the reduced state model, the control terms, and the measurement error. These statistics are imposed here in the form of diagonal covariance matrices estimated from the observations themselves. With more data, more realistic second-order statistics could be obtained using adaptive filter theory¹⁵.

The second difficulty is the existence of large GCM and data biases (primarily large-scale temperature offsets and missing wind-driven mixed-layer processes in the GCM, and geoid and atmospheric load errors in the altimetry) which need to be removed in the present approach. This is accommodated by decomposing the reduced state temperature, $T(\mathbf{r}, z, t)$, into a time-mean, $T_m(\mathbf{r}, z)$, a time-evolving, $T_e(z, t)$, and a residual term, $T'(\mathbf{r}, z, t)$:

$$T(\mathbf{r}, z, t) = T_m(\mathbf{r}, z) + T_e(z, t) + T'(\mathbf{r}, z, t) \quad (1)$$

where $\mathbf{r}$ is a horizontal position vector, z is depth, and t is time. $T_m(\mathbf{r}, t)$ is a correction for the large-scale offset of the GCM, and $T_e(z, t)$ is a correction for the inadequate GCM representation of seasonal warming in the mixed layer. The analysis is performed so that $T_m(\mathbf{r}, z)$ and $T_e(z, t)$ are rendered invisible in the data, thus eliminating the GCM/data biases, with the focus here being on estimating $T'(\mathbf{r}, z, t)$. The representation in equation (1) is neither unique nor rigorous, but it is adequate for determining the major modes of oceanic 'time variability'.

To test the accuracy of the constrained solution, the GCM/data combination is effected in two stages. In the first stage, the tomography data are withheld, and the GCM/altimetric combination is used to predict the spatial anomalies of the tomographic data (Fig. 3). The tomography shows the gradual heating of tomographic section H-W3 relative to the other two from February to October 1994. The GCM alone (without the altimetric data) fails to predict this relative warming trend, but the GCM/altimetric combination is consistent with the tomographic integrals.

In the second stage, both the altimetric and the tomographic data are used (Fig. 3). The addition of the tomographic data does not

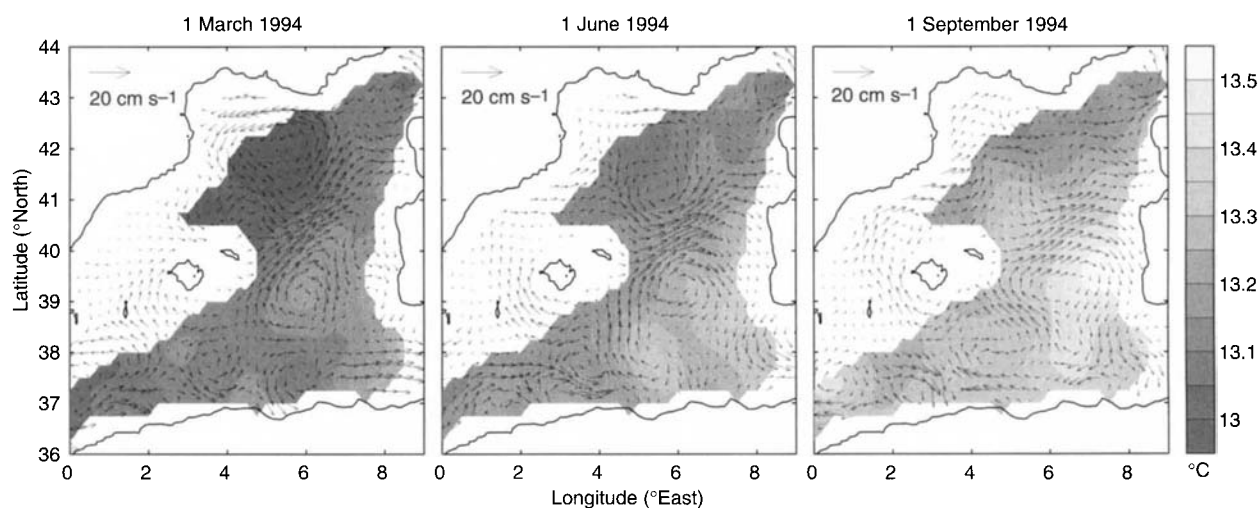


Figure 4 Estimates of western Mediterranean circulation, on 1 March, 1 June and 1 September, 1994, obtained by combining observations and the GCM. The grey-scale shading indicates depth-averaged (0–2,000 m) potential temperature and the arrows indicate horizontal velocity at 40 m depth. To summarize the key contributions of data and GCM, hydrographic data are used to estimate the time-mean temperature, the tomographic integrals provide the time-evolving temperature, the altimetry provides horizontal resolution, and the GCM determines the effect of observed winds and provides a reduced-state linear model which is used to interpolate the data to regions without measurements. The resulting estimates reproduce known features of the western Mediterranean circulation. For example

the coldest waters are detected in the Gulf of Lions, a region where deep convection is known to occur every year²², and are associated with the existence of a strong cyclonic gyre on 1 March, which gradually subsides later in the year. The warmest temperatures are observed in the southeastern part of the domain where the Levantine Intermediate Water enters the Algero-Provençal basin²³. The present estimates are preliminary and their detailed analysis and interpretation is beyond the scope of this Letter, nevertheless these estimates illustrate the combined power of ocean acoustic tomography, satellite altimetry and GCMs for oceanographic studies.

significantly change the estimates of the residual temperature $T'(\mathbf{r}, z, t)$, but it does reduce the uncertainty.

The remaining contributions to the Mediterranean circulation, $T_m(\mathbf{r}, z)$ and $T_c(z, t)$ in equation (1), are estimated without the assistance of the GCM. The tomographic data provides estimates of the time-evolving component $T_c(z, t)$, and a climatological analysis of hydrographic data¹⁰ provides an estimate of the time-mean component $T_m(\mathbf{r}, z)$. Last, the circulation is recovered using $T(\mathbf{r}, z, t)$ and the climatological salinity to compute density perturbations, and to obtain geostrophic corrections for the GCM circulation (the depth-integrated GCM transport is unchanged). Thus, to within the error bars and for the scales and processes resolved by the reduced state and by the climatologies, we obtain a complete description of the time-evolving three-dimensional Mediterranean circulation (Fig. 4).

Notwithstanding shortcomings in the existing models and data, the present results indicate that useful estimates of basin-scale oceanic circulation are already possible. Model and data deficiencies will diminish in time, but will always exist. The major issue is that the statement about model and data errors must be quantitative. Our results can be refined in several ways. In particular, a higher-resolution GCM with more realistic surface forcing, mixed-layer physical processes, and eddy parametrization will reduce GCM/data discrepancies. The estimates will also be improved when a more complete analysis of the THETIS-2 tomographic data becomes available, that is, more sections and higher vertical resolution. The estimation procedure itself can be augmented to account for smaller scales and more complete physical processes^{16,17}, to use better *a priori* statistics, or to obtain joint estimates of the three terms in equation (1).

The advent of high-accuracy altimeter missions and the improvement in GCM capability¹⁸ have brought oceanographers nearer to the goal of a true global observing system. What is still missing is the continuous coverage within the oceanic water column necessary to complement the global horizontal observations from space. The Acoustic Thermometry of Ocean Climate¹⁹ (ATOC) observations, now underway in the Pacific Ocean, in conjunction with the altimetry and developing models, will over the next several years produce a full oceanic-basin version of just such a system. The type of estimates we have made for the western Mediterranean Sea are now feasible for the entire North Pacific Ocean. □

Received 18 June; accepted 26 November 1996.

- Munk, W. & Wunsch, C. *Phil. Trans. R. Soc. Lond. A* **307**, 439–464 (1982).
- Send, U. *et al. Nature* **385**, 615–617 (1997).
- Fu, L.-L. *et al. J. Geophys. Res.* **99**, 24369–24381 (1994).
- Wunsch, C. & Stammer, D. *J. Geophys. Res.* **100**, 24895–24910 (1995).
- Munk, W., Worcester, P. & Wunsch, C. *Ocean Acoustic Tomography* (Cambridge Univ. Press, New York, 1995).
- Larnicol, G., Le Traon, P.-Y., Ayoub, N. & De Mey, P. *J. Geophys. Res.* **100**, 25163–25177 (1995).
- Stammer, D. & Wunsch, C. *J. Geophys. Res.* **99**, 24584–24604 (1994).
- Marshall, J., Hill, C., Perelman, L. & Adcroft, A. *J. Geophys. Res.* (in the press).
- Marshall, J., Adcroft, A., Hill, C., Perelman, L. & Heisey, C. *J. Geophys. Res.* (in the press).
- Brasseur, P., Beckers, J. M., Brankart, J. M. & Schoenauen *Deep-Sea Res.* **43**, 159–192 (1996).
- Zavatarelli, M. & Mellor, G. L. *J. Phys. Oceanogr.* **25**, 1384–1414 (1995).
- Roussenov, V., Stanev, E., Artale, V. & Pinardi, N. *J. Geophys. Res.* **100**, 13515–13538 (1995).
- Menemenlis, D. & Wunsch, C. *J. Atmos. Oceanic Technol.* (in the press).
- Rauch, H. E., Tung, F. & Striebel, C. T. *Ann. Inst. Astronaut. Astronaut.* **3**, 1445–1450 (1965).
- Blanchet, I., Frankignoul, C. & Cane, M. A. *Mon. Weath. Rev.* **125**, 40–58 (1997).
- Marotzke, J. & Wunsch, C. *J. Geophys. Res.* **98**, 20149–20167 (1993).
- Schiller, A. *J. Mar. Res.* **53**, 453–497 (1995).
- Stammer, D., Tokmakian, R., Semtner, A. & Wunsch, C. *J. Geophys. Res.* **101**, 25779–25811 (1996).
- Forbes, A. *Sea Technol.* **35**, 65–67 (1994).
- Fiegluth, P. W., Menemenlis, D., Ho, T., Willsky, A. & Wunsch, C. *J. Atmos. Oceanic Technol.* (submitted).
- Bryden, H. L., Candela, J. & Kinder, T. H. *Prog. Oceanogr.* **33**, 201–248 (1994).
- Send, U., Schott, F., Gaillard, F. & Desaubies, Y. *J. Geophys. Res.* **100**, 6927–6941 (1995).
- Tziperman, E. & Malanotte-Rizzoli, P. *J. Mar. Res.* **49**, 411–434 (1991).

Acknowledgements. We thank our colleagues D. Stammer, C. King and P. Fiegluth for help with the analysis of the TOPEX/POSEIDON altimeter data, and J. Marshall, C. Heisey and A. Adcroft for help with the GCM. This work was supported by the Strategic Environmental Research and Development Program and the Advanced Research Projects agency (ARPA) as part of the ATOC project; by NASA; and by grants of High Performance Computer time from Project SCOUT at the M.I.T. Laboratory for Computer Science, and from the Arctic Region Supercomputing Center.

Correspondence and requests for materials should be addressed to D.M. (e-mail: dimitri@gulf.mit.edu).

Rheology of the continental lithosphere inferred from sedimentary basins

Robert Newman & Nicky White

Department of Earth Sciences, Bullard Laboratories, University of Cambridge, Madingley Rise, Madingley Road, Cambridge CB3 0EZ, UK

The steady-state flow properties of the continental lithosphere play an important role in a wide range of geological processes¹. A complete dynamic description of lithospheric deformation requires information about the magnitude of driving forces and the rheology of the crust and lithospheric mantle, about which there is little agreement^{2–6}. Here we constrain these properties by analysing variations in strain rate during the extension of continental lithosphere. We determine the temporal variation of strain rate from the subsidence curves of a global sample of Phanerozoic sedimentary basins. The peak strain rate and final strain estimated from these strain-rate histories suggest that the cessation of extension is governed by cooling and concomitant strengthening of the underlying lithospheric mantle. Dynamic modelling of these data indicates that the rheology of the lithosphere is controlled by power-law creep with a stress exponent of three and an activation energy of $\sim 500 \text{ kJ mol}^{-1}$. This rheology is consistent with that inferred from laboratory experiments on dry olivine⁷ extrapolated to lithospheric conditions.

Kinematic models of lithospheric extension account satisfactorily for the structure and evolution of many sedimentary basins. But there is little agreement about the main aspects of the dynamical problem, namely the origin and magnitude of forces that drive extension and the rheology of the crust and lithospheric mantle^{2–6}. We also do not know what limits the value of the stretching factor, β , defined as the ratio between initial and final crustal thickness. One possibility is that variation in strain rate is controlled by changes in the driving force. Alternatively, lithospheric mantle may cool and strengthen as the Moho rises, thereby reducing the strain rate produced by a given force². If such cooling controls basin evolution, then we expect a characteristic change of strain rate with time⁵. Temporal strain-rate variation can therefore be used to test dynamic models⁵.

The strength of the lithosphere (that is, the stress required to deform it at a given strain rate) is usually estimated from laboratory experiments in which quartz and olivine are deformed under high pressure at strain rates of 10^{-7} – 10^{-4} s^{-1} and at temperatures of $\geq 1,300^\circ\text{C}$ (ref. 7). Results show that the rheological properties of quartz and olivine can be described by empirical constitutive

Table 1 Principal parameters used in numerical experiments

a	Lithospheric thickness	125 km
t_c	Crustal thickness	33 km
A	Pre-exponential constant	$1 \times 10^5 \text{ MPa}^{-3} \text{ s}^{-1}$
n	Stress exponent	3
Q	Activation energy	300–1,000 kJ mol^{-1}
R	Universal gas constant	$8.314 \text{ J mol}^{-1} \text{ K}^{-1}$
T_M	Initial Moho temperature	580°C
T_L	Asthenosphere temperature	$1,330^\circ\text{C}$
k_c	Thermal conductivity of crust	$2.5 \text{ W m}^{-1} \text{ }^\circ\text{C}^{-1}$
k_m	Thermal conductivity of mantle	$3.0 \text{ W m}^{-1} \text{ }^\circ\text{C}^{-1}$
H	Crustal heat production	$1.5 \mu\text{W m}^{-3}$
T	Absolute temperature	K
β	Uniform stretching factor	
τ	Deviatoric stress	Pa
ϵ	Strain rate	s^{-1}

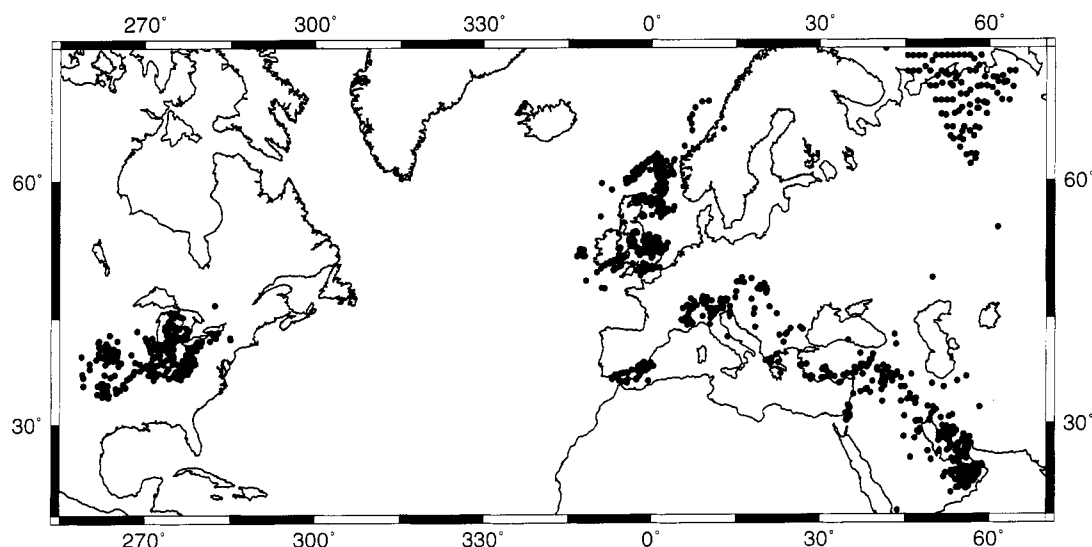


Figure 1 Mercator projection showing global distribution of stratigraphic sections which were derived from field logging, industry wells and seismic reflection data. Regions were chosen to include continuous, shallow-water marine, sections from different portions of the geological timescale. Past water depths are generally well constrained, and syn-rift topography has wavelengths of ~ 20 km with small amplitudes so that the flexural rigidity of the lithosphere can be ignored.

equations which relate strain rate to deviatoric stress, temperature and confining pressure^{8–10}. When the homologous temperature (ratio of temperature to melting temperature, both in K)⁸ is ~ 0.4 , the relationship between $\dot{\epsilon}$, the strain rate, and τ , the deviatoric stress, is given by

$$\dot{\epsilon} = A\tau^n \exp(-Q/RT) \quad (1)$$

where A , n and Q are constants that depend on the material¹⁰.

Application of laboratory results to the continental lithosphere depends on a range of important assumptions. Experimental strain rates are much faster than geological ones ($< 10^{-14} \text{ s}^{-1}$), samples are usually monomineralic not polyphase aggregates, and the effects of melt, hydrous phases and grain size are uncertain¹¹.

To determine strain-rate histories, we and our colleagues have collated and digitized over 2,000 stratigraphic sections from extensional sedimentary basins world-wide (Fig. 1). These data span the past 600 Myr and are from regions where stratigraphic sequences are as complete as possible. The lack of deeply penetrating boreholes means that data from the highly extended portions of continental margins are lacking. Tectonic subsidence has been calculated and modelled using an inversion algorithm to determine temporal strain-rate variation (Fig. 2)^{12–14}. Data inversion yields excellent fits between theoretical and observed subsidence, and strain-rate profiles are generally consistent with independent geological information about the duration and number of rift periods.

To compare these data with the results of numerical modelling, it is convenient to plot the stretching factor for each event against the peak strain rate (Fig. 3a). Note that there are often several episodes of extension within a basin and that each of these events has been plotted as a separate point. A rifting episode is assumed to have ceased if the strain rate falls below 10^{-17} s^{-1} . At low peak strain rates ($< 10^{-16.5} \text{ s}^{-1}$), total stretching factors are also low (< 1.03) while higher strain rates give rise to higher total stretching factors. It is important to point out that the same systematic relationship and range of strain rates are observed when data from individual basins are plotted. This pattern is difficult to explain if lithospheric extension is halted by the simple removal of driving forces, which would produce the same correlation in each basin only if such forces always last about 25 Myr. If, on the other hand, extension is driven by buoyancy forces¹⁵, the force does decrease as strain increases and could limit extension without strengthening the lithosphere. However, to produce the same correlation, each basin would need the same initial strength.

Our dynamic model determines the strain rate and temperature history at each point on a two-dimensional cross-section through the lithosphere, which is treated as a continuum. The equations of motion and of thermal diffusion are solved by a finite-element

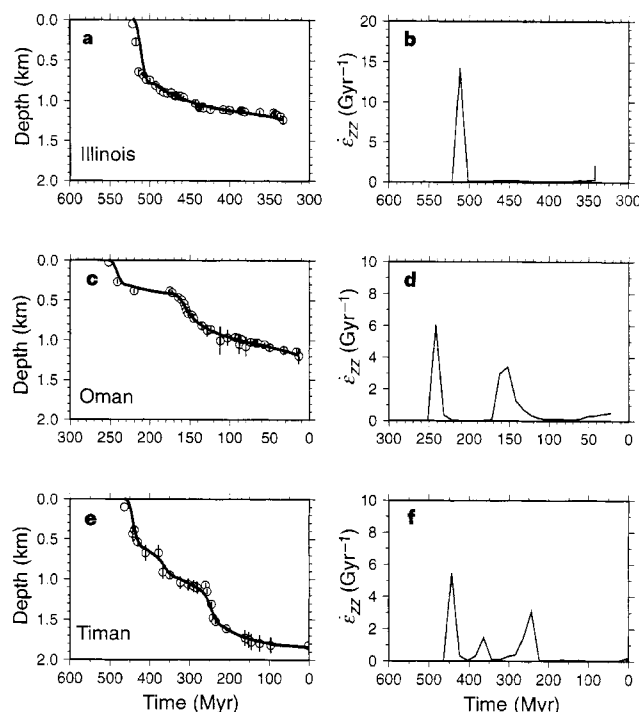


Figure 2 Strain rate inversion of three stratigraphic sections which illustrate the variety of subsidence data used. **a, c, e**, Tectonic (that is, water-loaded) subsidence data (circles) produced by standard decompaction and back-stripping; vertical bars indicate variance; solid line, best-fitting theoretical subsidence curve. **b, d, f**, Linearly interpolated strain-rate distribution obtained by inversion and corresponding to best-fitting subsidence curve. No *a priori* information about the duration of rifting, the number of rift episodes, or the total stretching factor is required. Error analysis demonstrates that strain-rate peaks are resolved to better than a factor of two (95% confidence level)¹⁶, reflecting structure within the observed subsidence data. Total stretching factors (β) are: Illinois, 1.18; Oman, 1.19; Timan, 1.31. Multiply by 3.17×10^{-17} to convert units from Gyr^{-1} to s^{-1} .

method¹⁶. Temperature changes in the crust cannot halt extension because the depth-integrated crustal strength does not increase above its initial value. Therefore, in order to isolate the effect of lithospheric mantle cooling, the crust is assigned a low viscosity which minimizes its effect on integrated lithospheric strength (we used $Q = 0 \text{ kJ mol}^{-1}$, $n = 3$ and $A = 10^{-15} \text{ MPa}^{-3} \text{ s}^{-1}$). A set of experimental results are generated for a given rheology by varying the driving force, which is applied on one side boundary. Varying the initial Moho temperature, T_M , instead of the force yields similar solutions. Principal parameters are given in Table 1.

The results shown in Fig. 3a suggest that an activation energy, Q ,

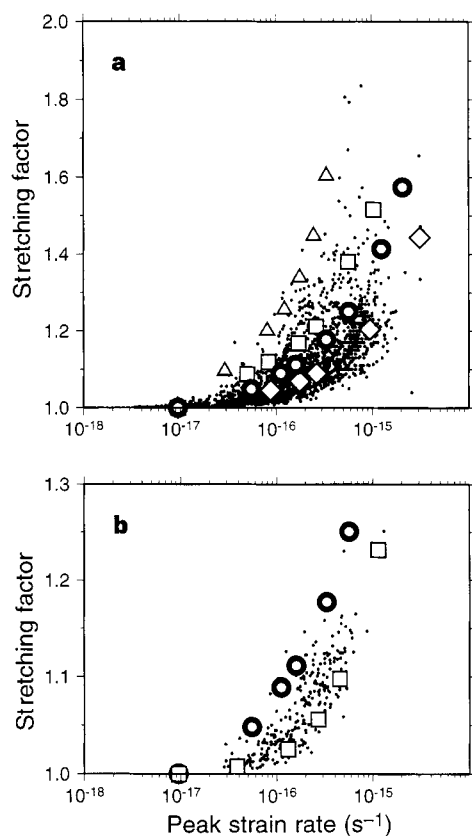


Figure 3 a, Small dots, lithospheric stretching factor (β) plotted as a function of peak strain rate for 2,148 separate rift episodes. Note that each of the separate rift episodes shown in Fig. 2b and c plot as individual points. Data show that continental lithosphere away from highly extended passive margins stretches at peak strain rates of 10^{-17} – 10^{-15} s^{-1} . Strain rates have errors of a factor of two, stretching factors have errors of ~ 0.05 . Big symbols represent four different sets of results obtained from numerical experiments. Each symbol type represents one series of numerical experiments carried out for a given set of rheological parameters. In each case, the cumulative stretching factor reached once strain rate has dropped to 10^{-17} s^{-1} is plotted against peak strain rate. Open squares, $Q = 300 \text{ kJ mol}^{-1}$; circles, $Q = 500 \text{ kJ mol}^{-1}$; open diamonds, $Q = 1,000 \text{ kJ mol}^{-1}$. Values of other parameters are given in Table 1. Open triangles, set of numerical experiments which include loading and thermal effects of basin sedimentation; sediment density, 2.1 Mg m^{-3} ; thermal conductivity of sediments varies as a function of depth between 1.1 and $2.1 \text{ W m}^{-1} \text{ }^{\circ}\text{C}^{-1}$; $Q = 500 \text{ kJ mol}^{-1}$. **b**, Small dots, sub-set of data from North America and the North Sea which only include second rift events. Solid circles, $Q = 500 \text{ kJ mol}^{-1}$ as in **a**; open squares, $Q = 500 \text{ kJ mol}^{-1}$ but temperature structure affected by a previous rift event. Previous β factor = 1.4 , second rift event occurs 50 Myr later. In **a** and **b**, governing equations are solved using a finite-element mesh made up of six-node triangular elements. The driving force is constant and applied at one side boundary. There is no horizontal variation of rheology so the lithosphere deforms by pure shear. Values of key parameters are given in Table 1, further details are given in ref. 16.

of $\sim 500 \text{ kJ mol}^{-1}$ results in a good fit to the global dataset. During extension, the temperature structure of the lithosphere is perturbed and the temperature at any material point starts to decrease by thermal diffusion¹². To lower the strain rate by an order of magnitude, the temperature of material at an initial temperature of T_0 must be lowered by about $2.3RT_0^2/Q$ (from equation (1)). If the strongest part of the lithosphere is just below the Moho and if Q is 500 kJ mol^{-1} , then the temperature drop is sufficient to halt stretching after 20 – 30 Myr . A lower value of Q requires a larger temperature drop and permits a longer period of stretching, thus producing a curve with a steeper gradient (Fig. 3a).

Thinning the lithosphere concentrates the same force over less depth and causes a rise in strain rate early in the stretching period⁵. This effect is larger for higher values of n , the stress exponent, and for higher strain rates. At lower strain rates ($< 10^{-15.5} \text{ s}^{-1}$), $Q = 500 \text{ kJ mol}^{-1}$ fits the data irrespective of the value of n , whereas at higher strain rates, $n = 3$ yields the best fit. Changing A , the pre-exponential constant, has the same effect as changing the driving force and does not affect the correlation.

The three most important complications which could affect our conclusions are: (1) basin sedimentation¹⁷; (2) a crustal layer which is strong compared with the lithospheric mantle; and (3) brittle deformation of the lithospheric mantle⁴. A low-conductivity sedimentary infill reduces the temperature perturbation generated during lithospheric stretching because the equilibrium geotherm changes as sediment is deposited. For a given strain rate, sedimentary infill may cause β to increase significantly (see triangles in Fig. 3a) but cooling of the lithospheric mantle still limits the total strain. Results are very sensitive to sedimentation rate, to changes in thermal conductivity and diffusivity caused by compaction of the sediment pile, and to any original surface elevation above sea level.

If the crust is initially stronger than the lithospheric mantle, the required temperature drop, and thus β , will increase. For example, the Moho temperature might drop by 20°C before crust and lithospheric mantle have equal strength, and by a further 40°C to halt extension. It is therefore difficult to achieve small β factors if the crust is strong.

As the lithospheric mantle cools, it may enter the brittle deformation regime. This effect will limit the strength increase at a given depth but the overall strength of the lithosphere can still increase¹⁸. In such cases, the peak strain rate will occur later in the stretching period.

We suggest that sedimentary and crustal effects account for the scatter above the $Q = 500 \text{ kJ mol}^{-1}$ set of numerical experiments (Fig. 3a). Many data points which fall below the $Q = 500 \text{ kJ mol}^{-1}$ line are not from first rift episodes but are from second and third events which occur 40 – 100 Myr later. A significant thermal perturbation can persist for up to 120 Myr after an event and such a perturbation will give rise to faster cooling (that is, smaller β s) during later rifting periods. A subset of second rifting events fall nicely between the standard solution for $Q = 500 \text{ kJ mol}^{-1}$ and a set of solutions which incorporate the thermal perturbation caused by a prior rift episode (Fig. 3b). Minor temporal variations of the driving force and its modification by buoyancy forces may be responsible for some of the scatter.

Four principal features of the global dataset are supported by our cooling hypothesis: (1) strain rates from basins are slow enough to allow significant thermal diffusion; (2) the correlation between peak strain rate and final stretching factor is similar for all examined basins; (3) the power-law rheology required to fit these data is consistent with results of laboratory experiments; and (4) episodes which occur less than $\sim 120 \text{ Myr}$ after previous events show the predicted effect on β . These features are difficult to explain by removal of driving forces alone. We estimate that the rheology of the continental lithosphere is best modelled by equation (1) with $Q = 500 \text{ kJ mol}^{-1}$ and $n = 3$. These values suggest that lithospheric rheology is controlled by power-law creep of dry olivine, vindicating

the use of laboratory results in estimating flow properties of the lithosphere. □

Received 25 July 1996; accepted 2 January 1997.

1. Ranalli, G. *Rheology of the Earth* (Chapman & Hall, London, 1995).
2. England, P. J. *Geophys. Res.* **88**, 1145–1152 (1983).
3. Kuznir, N. J. & Park, R. G. *Geophys. J. R. Astron. Soc.* **79**, 513–538 (1984).
4. Sawyer, D. S. *J. Geophys. Res.* **90**, 3021–3025 (1985).
5. Sonder, L. J. & England, P. C. *J. Geophys. Res.* **94**, 7603–7619 (1989).
6. Bassi, G. *Geophys. J. Int.* **122**, 195–210 (1995).
7. Kohlstedt, D. L. & Goetze, C. J. *Geophys. Res.* **79**, 2045–2051 (1974).
8. Ashby, M. F. & Verrall, R. A. *Phil. Trans. R. Soc. Lond. A* **288**, 59–95 (1977).
9. Weertman, J. *Rev. Geophys. Space Phys.* **8**, 145–168 (1977).
10. Kirby, S. H. & Kronenberg, A. K. *Rev. Geophys.* **25**, 1219–1244 (1987).
11. Paterson, M. S. *Tectonophysics* **133**, 33–43 (1987).
12. Jarvis, G. T. & McKenzie, D. *Earth Planet. Sci. Lett.* **48**, 42–52 (1980).
13. White, N. *Nature* **366**, 449–452 (1993).
14. White, N. *Earth Planet. Sci. Lett.* **122**, 351–371 (1994).
15. Jones, C. H., Unruh, J. R. & Sonder, L. J. *Nature* **381**, 37–41 (1996).
16. Newman, R. B. thesis, Oxford Univ. (1992).
17. McKenzie, D. *Earth Planet. Sci. Lett.* **55**, 87–98 (1981).
18. Houseman, G. A. & England, P. C. *J. Geophys. Res.* **91**, 719–729 (1986).

Acknowledgements. We thank A. Butler, N. O. Leary and C. Trowell for permitting us to use their unpublished data in Fig. 2; S. Baldwin, P. Bellingham, J. Brodie, B. Hall, S. Lewis, D. Lyness, E. Rowley and D. Wooler for their help; and D. McKenzie, P. England and L. Sonder for comments. R.N. was supported by an NERC fellowship.

Correspondence should be addressed to N.W. (e-mail: nwhite@esc.cam.ac.uk).

Orbital forcing of deep-sea benthic species diversity

T. M. Cronin* & M. E. Raymo†

* Climate History Team, 955, US Geological Survey, Reston, Virginia 20192, USA

† Department of Earth, Atmosphere and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Explanations for the temporal and spatial patterns of species biodiversity focus on stability–time^{1–3}, disturbance–mosaic (biogenic microhabitat heterogeneity)^{4,5} and competition–predation (biotic interactions)^{6,7} hypotheses. The stability–time hypothesis holds that high species diversity in the deep sea and in the tropics reflects long-term climatic stability³. But the influence of climate change on deep-sea diversity has not been studied and recent evidence suggests that deep-sea environments undergo changes in climatically driven temperature⁸ and flux of nutrients⁹ and organic-carbon¹⁰ during glacial–interglacial cycles. Here we show that Pliocene (2.85–2.40 Myr) deep-sea North Atlantic benthic ostracod (Crustacea) species diversity is related to solar insolation changes caused by 41,000-yr cycles of Earth's obliquity (tilt). Temporal changes in diversity, as measured by the Shannon–Weiner index, $H(S)$, correlate with independent climate indicators of benthic foraminiferal oxygen-isotope ratios (mainly ice volume^{11–13}) and ostracod Mg:Ca ratios (bottom-water temperature⁸). During glacial periods, $H(S) = 0.2–0.6$, whereas during interglacials, $H(S) = 1.2–1.6$, which is three to four times as high. The control of deep-sea benthic diversity by cyclic climate change at timescales of $10^3–10^4$ yr does not support the stability–time hypothesis because it shows that the deep sea is a temporally dynamic environment. Diversity oscillations reflect large-scale response of the benthic community to climatically driven changes in either thermohaline circulation, bottom temperature (or temperature-related factors) and food, and a coupling of benthic diversity to surface productivity.

We examined ostracods from 125 10-ml samples from the deep-sea drilling project (DSDP) site 607, located on the western flank of the mid-Atlantic ridge (41° N, 19° W; water depth, 3,427 m), and DSDP site 610 near the Rockall plateau (53° N, 19° W; water depth, 2,417 m). Site 607 is located in the path of and is sensitive to deep water forming at high latitudes of the Nordic and Labrador Seas¹⁵. The study covers the period 2.85–2.40 Myr, a Pliocene interval when

continental ice sheets began to wax and wane at the 41,000-year obliquity period of orbital insolation¹⁵. Well documented climate cycles observed globally in ice volume, sea surface temperature and ocean chemistry signify a linear response of Earth's climate to solar insolation forcing^{11–16}, and provide an ideal series of natural experiments to test models that call for climatic influence on diversity.

Ostracods are the only metazoan taxon regularly fossilized in deep-sea sediments that allow quantitative analysis of faunal patterns¹⁷. Moreover, deep-sea ostracod species have distinct habitat preferences¹⁸, making the ostracod assemblage a subset of the deep-sea benthic community containing representative 'functional' groups¹⁹. Ostracods occur in variable numbers^{20,21}, averaging 60 specimens per sample (in the range 20–120). The number of species, S , varied from 2 to 15, but this is an underestimate because *Krithe* species were lumped (juveniles are difficult to differentiate; there were usually 2–5 species of *Krithe* per sample). We used the Shannon–Wiener function $H(S) = -\sum_i S p_i \ln(p_i)$ where $H(S)$ is diversity, and p_i is the proportion of species i , which is commonly used to measure diversity²² and is not as sensitive to abundance¹.

$H(S)$ is plotted against two measures of climate and oceanographic change obtained from site 607 for eleven obliquity cycles (Fig. 1). Oxygen isotope ratios of benthic foraminifers reflect mainly changes in global ice volume (about two-thirds of the total $\delta^{18}\text{O}$ signal) and bottom-water temperature changes (about one-third of the signal). Late Pliocene, 41,000-yr, glacial–interglacial $\delta^{18}\text{O}$ fluctuations are lower in amplitude than Late Quaternary 100,000-yr cycles^{13,23}. Changes in Mg:Ca ratios in *Krithe* signify mainly bottom-water temperature (BWT) change due to variations in the relative amount of cold Antarctic deep water (AABW) and warmer North Atlantic deep water (NADW) over the site. Glacial–interglacial changes in Mg:Ca averaged about $2.3 \text{ mmol mol}^{-1}$ for the 11 cycles, which equates to an average temperature change of 2.3°C . Glacial bottom-water temperatures were $\sim 1.0–1.5^\circ\text{C}$; those of interglacials were $\sim 3–4^\circ\text{C}$ (ref. 8).

Figure 1 shows a clear relationship between climate change and species diversity. Low diversity generally corresponds to isotopically heavy $\delta^{18}\text{O}$ values (large ice volume) and glacial climates; high diversity corresponds to light $\delta^{18}\text{O}$ (smaller ice volume) and interglacial climates. This relationship is expressed graphically in Fig. 2a ($r = -0.722$). All 11 of the isotopically defined glacial–interglacial cycles are expressed in the $H(S)$ measure, and only where there are missing isotope values (near 2.54, 2.62 and 2.65 Myr) are there slight offsets between the two curves. Even more striking is that the amplitude of the $H(S)$ measure is greatest during the strongest cycles; $H(S)$ drops the most during the strongest glacial events (for example, those at 2.70, 2.50, 2.46 and 2.42 Myr). These last three glacial events, also characterized by a pronounced flux of ice-rafted detritus to this site, are referred to as isotopic stages 100, 98 and 96, and can be identified in many oceanic proxy records¹³. To investigate further the leads and lags of diversity with environmental indicators, cross-spectral analysis¹⁴ and a new timescale²⁴ show that $H(S)$ maxima lag behind the maxima in Mg:Ca ratios (warmest bottom-water temperatures) by 1.6 ± 1.5 kyr and lead ahead of the minima in $\delta^{18}\text{O}$ (maximum interglacial conditions) by 2.8 ± 1.7 kyr. The correlation between $H(S)$ and Mg:Ca is shown in Fig. 2b ($r = 0.53$). Benthic species diversity is highest just before the minimum ice volume and least just before the maximum ice volume. Figure 2c shows no correlation between abundance of ostracods and oxygen isotope values, suggesting that climatically induced environmental changes do not appear to control abundance. Given the importance of broad spatial and temporal scales in all studies of biodiversity, we investigated whether the patterns observed at site 607 occur in correlative Pliocene deep-sea sediments from other North Atlantic cores or during younger climatic cycles in the same region. Data from the same 2.85–2.50-Myr interval at DSDP site 610 near the Rockall

plateau also show oscillations in $H(S)$ values of 0.2–0.6 during glacials and 1.3–1.6 during interglacials (Fig. 3a); thus, similar diversity changes occurred in the eastern North Atlantic without any obvious relation between $H(S)$ and abundance (Fig. 3b). In Chain core 82–24 near site 607, ostracode diversity over the past 150,000 years (including two glacial cycles) shows high diversity in interglacial periods and glacial periods that are depauperate.

It is important to exclude other factors that may cause apparent secular diversity changes. Carbonate dissolution can selectively affect shells of surface-dwelling planktic foraminifers after they settle to the bottom, especially during Late Quaternary glacial periods when the lysocline in the Atlantic shoaled several hundred metres from its current 4,000 m depth. However, dissolution is not considered to be the cause of the ostracod diversity patterns for several reasons. Deep-sea ostracods are adapted to deep-sea environments, and ecological and zoogeographical factors, not post-mortem alteration, explain their modern and fossil deep-sea distribution^{17,18,20,21,25}. Ostracods secrete a carapace formed by an interlocking calcitic–chitinous lattice, with a thin outer organic epicuticle that helps protect the shell from dissolution; the carapace is thicker and less porous than shells of planktic foraminifers, which are adapted to ocean surface layers. Moreover, Pliocene glacials were only half the amplitude of Late Quaternary glacials, so it is doubtful whether lysocline shoaling would impact ostracods at 3,400 m and 2,400 m at sites 607 and 610. Nor would we expect to see similar diversity patterns in cores from both western and eastern North Atlantic basins, if bottom-water changes were selectively influencing species preservation, especially as site 610 is too shallow and isolated from the influence of potentially corrosive AABW. Other evidence argues against dissolution: occurrence in all samples of *Krithe*

onments, and ecological and zoogeographical factors, not post-mortem alteration, explain their modern and fossil deep-sea distribution^{17,18,20,21,25}. Ostracods secrete a carapace formed by an interlocking calcitic–chitinous lattice, with a thin outer organic epicuticle that helps protect the shell from dissolution; the carapace is thicker and less porous than shells of planktic foraminifers, which are adapted to ocean surface layers. Moreover, Pliocene glacials were only half the amplitude of Late Quaternary glacials, so it is doubtful whether lysocline shoaling would impact ostracods at 3,400 m and 2,400 m at sites 607 and 610. Nor would we expect to see similar diversity patterns in cores from both western and eastern North Atlantic basins, if bottom-water changes were selectively influencing species preservation, especially as site 610 is too shallow and isolated from the influence of potentially corrosive AABW. Other evidence argues against dissolution: occurrence in all samples of *Krithe*

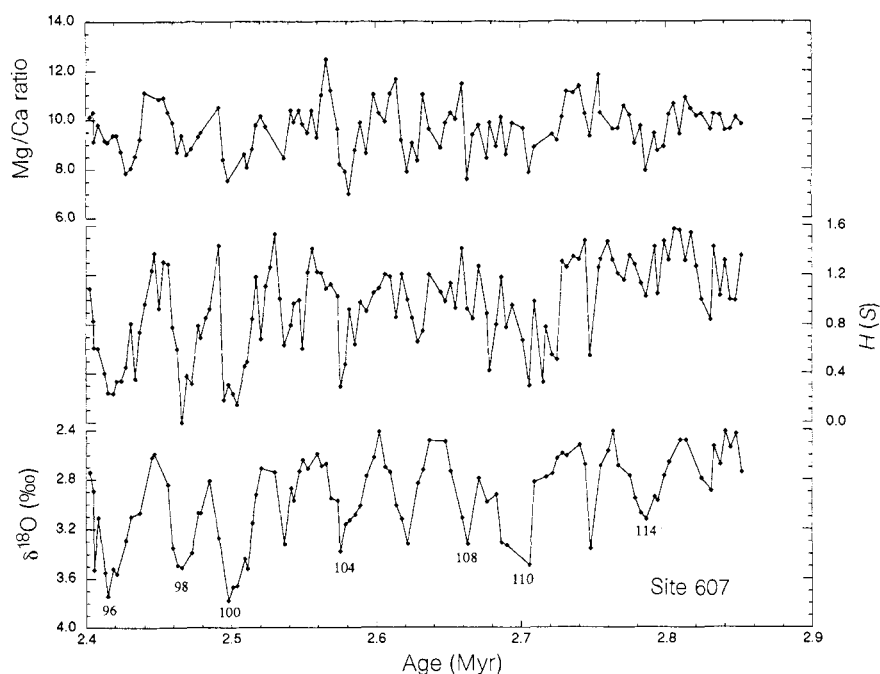


Figure 1 Deep-sea ostracod species diversity at DSDP site 607 measured by the Shannon-Wiener information function $H(S)$; $H(S)$ against benthic foraminiferal $^{18}\text{O}/^{16}\text{O}$ ratios^{12,13} (mainly global ice volume); heavier-isotope values reflect high ice volume during glacial periods; lighter-isotope values signify interglacial warm periods with low ice volume. Mg:Ca ratios (from ref. 9) are indicative of bottom-

water temperatures: lower ratios indicate that temperatures were less than about 1 to 2°C; higher ratios signify warmer temperatures of ~3–4°C. The age model for site 607 (ref. 12) gives a sampling interval of ~3,500–4,000 yr. Selected oxygen isotope stages are indicated.

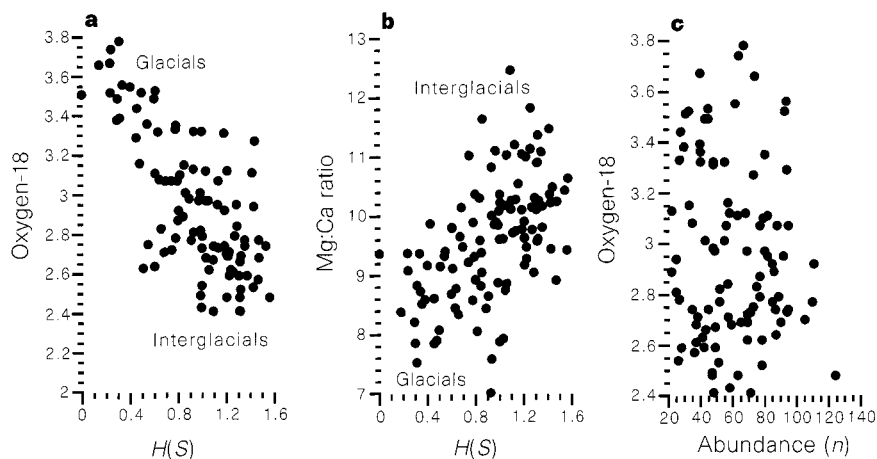


Figure 2 **a**, Relation between $H(S)$ and benthic foraminiferal isotope values ($r = -0.72$); **b**, relation between $H(S)$ and Mg:Ca ratios ($r = 0.53$); **c**, relation between the abundance n and oxygen isotope values.

juveniles (which have high Mg:Ca ratios and are thus more susceptible to dissolution), absence of thick-shelled, dissolution-resistant genera (*Henryhowella*) in glacial intervals, 'asymmetry' of faunal assemblages during preglacial and deglacial phases of each cycle²⁶, and lack of correlation between either valve fragmentation or species proportions and foraminifer $\delta^{13}\text{C}$, an indicator of AABW.

Whereas seasonal and geographical changes in the distribution of solar insolation caused by variations in the Earth's tilt drive deep-sea benthic diversity, the translation of orbitally induced changes to the deep-sea ecosystem must involve complex community interactions that occur during each cycle. These may involve the creation of new niches and/or niche partitioning which reflect community restructuring in response to environmental change. For instance, a bottom-water temperature change of 2–3 °C is small relative to those occurring in mid-high latitude, shallow marine environments during climatic cycles, but it is nonetheless large compared to short-term deep-sea temperature changes. Temperature itself may be a direct factor responsible for diversity changes, but as $H(S)$ lags behind BWT by only ~1,500 yr, BWT change might catalyse a series of biotic interactions that take 1,500 yr to equilibrate.

A second environmental change that could have an impact on diversity involves nutrients (for example, phosphorus and nitrogen). Modern NADW is relatively low in nutrients; AABW is high in nutrients. Cadmium/calcium ratios in benthic foraminifera (proxy for phosphorus) and $\delta^{13}\text{C}$ oscillations have been used to trace changes in bottom-water circulation and infer nutrient changes^{9,13}. At site 607, the $\delta^{13}\text{C}$ record shows that Pliocene glacial periods of low diversity correspond to inferred periods of high nutrient concentrations and vice versa¹³. Thus inferred nutrient content of bottom waters is inversely correlated with diversity, and high nutrient content does not lead to higher diversity.

A third, more plausible explanation of diversity involves changing food resources stemming from surface-water productivity changes. Climatically induced surface productivity changes are coupled to benthic organisms²⁷ and occur during the last deglaciation in benthic foraminiferal assemblages¹⁰. Dinoflagellate changes in late Pliocene sediments at site 607 record major changes in surface water

productivity²⁷, also indicated by planktic foraminifers and discoasters cycles²⁸, which would lead to decreased food resources for benthic organisms like ostracodes. Thus, long-term diversity changes in the deep-sea benthos are probably reflecting changing total abundance and/or temporal variability in food resources^{5,10}, and their effect on the community, a factor that also explains large-scale spatial diversity patterns²⁹. Decreased or more variable food supply during glaciations may increase stress in the deep-sea community, increasing competition for limited resources and possibly also decreasing habitat (sediment?) heterogeneity on small spatial scales.

Where then do those taxa that inhabit the deep North Atlantic during interglacials live during glacial periods? Faunal evidence suggests that during the last glacial some deep-sea ostracods (*Henryhowella*, *Bradleya*, some *Krithe*) inhabited mid-depth (1,000–2,000 m) bathyal environments in the Arctic Ocean³⁰ and off the Bahama islands. Bathyal regions may provide the resource requirements and/or physical chemical conditions necessary to survive and may therefore act as an ocean-margin refuge from which species can later return to deep-sea abyssal environments following deglaciation. An analogous situation exists when shallow marine organisms show poleward zoogeographic range expansion of thermophilic species during interglaciations and equatorward expansion of cryophilic species during glaciations.

The correlation of deep-sea benthic species diversity with orbitally induced cyclic changes challenges the stability-time hypothesis³, which holds that patterns of high species diversity in the deep sea result from long-term environmental stability that allows new species to evolve over time. Deep-sea environments are dynamic in temperature and food over timescales of 10^3 – 10^5 yr; our results show that biodiversity is also inherently unstable. Thus, assumptions of constant deep-sea resources do not hold over timescales of 10^3 – 10^5 yr. Likewise, temporal patterns of diversity do not appear to be random, as are some diversity patterns over small spatial scales (centimetres to metres)⁵, but instead are linked to external, abiotic solar insolation changes. Our data also underscore the importance of historical factors in understanding biodiversity over long timescales and on regional and ocean basin scales, and suggest that predicting the impact of anthropogenically forced climate change on diversity will remain difficult until we gain a better understanding of how climate has affected biodiversity in the past.

Received 4 May; accepted 11 December 1996.

1. Sanders, H. L. *Am. Nat.* **102**, 243–282 (1968).
2. Fischer, A. G. *Evolution* **14**, 64–81 (1960).
3. Hessler, R. R. & Sanders, H. L. *Deep-Sea Res.* **14**, 65–78 (1967).
4. Hutchinson, G. E. *Am. Nat.* **93**, 145–159 (1959).
5. Grassle, J. F. & Maciolek, N. J. *Am. Nat.* **139**, 313–341 (1992).
6. Paine, R. T. *Am. Nat.* **100**, 65–75 (1966).
7. Pianka, E. M. *Am. Nat.* **100**, 33–46 (1966).
8. Dwyer, G. D. *et al. Science* **270**, 1347–1351 (1995).
9. Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784–787 (1982).
10. Thomas, E., Booth, L., Maslin, M. & Shackleton, N. J. *Paleoceanography* **10**, 545–562 (1995).
11. Shackleton, N. J. & Opdyke, N. *Quat. Res.* **3**, 39–55 (1973).
12. Raymo, M. E., Ruddiman, W. F., Backman, J., Clement, B. M. & Martinson, D. G. *Paleoceanography* **4**, 413–446 (1989).
13. Raymo, M. E., Hodell, D. & Jansen, E. *Paleoceanography* **7**, 645–672 (1992).
14. Imbrie, J. *et al. Paleoceanography* **7**, 701–738 (1992).
15. Shackleton, N. J. *et al. Nature* **307**, 620–623 (1984).
16. Berger, A. L. *Quat. Sci. Rev.* **11**, 571–582 (1992).
17. Whatley, R. *Application of Ostracoda* (ed. Maddocks, R. F.) 51–77 (University of Houston Press, Texas, 1983).
18. Benson, R. H. *Paleogeogr. Palaeoclimatol., Palaeoecol.* **48**, 107–141 (1984).
19. Huston, M. A. *Biological Diversity* (Cambridge University Press, Cambridge, 1994).
20. Whatley, R. C. & Coles, G. *Revista Espanola de Micropaleontologia* **19**, 33–97 (1987).
21. Cronin, T. M. *Microfossils and Oceanic Environments* (eds Moquilevsk, A. & Whatley, R. C.), (University of Wales, Aberystwyth, 1996).
22. Buzas, M. A. *Taxon* **21**, 275–286 (1972).
23. Imbrie, J. *et al. Paleoceanography* **8**, 699–736 (1993).
24. Berggren, W. A., Kent, D. V., Swisher, C. III & Aubrey, M.-P. *SEPM Spec. Publ. No. 54*, 129–212 (1995).
25. Dingle, R. V. & Lord, A. R. *Paleogeogr., Palaeoclimatol., Palaeoecol.* **99**, 213–235 (1990).
26. Cronin, T. M., Raymo, M. E. & Kyle, K. P. *Geology* **24**, 695–698 (1996).
27. Versteeg, G. *Mar. Micropal.* **23**, 147–183 (1994).
28. Backman, J., Pestioux, P., Zimmerman, H. & Hermelin, O. *North Atlantic Paleoceneography* (Geol. Soc. Amer. Spec. Publ. No. 21) 231–242 (1986).

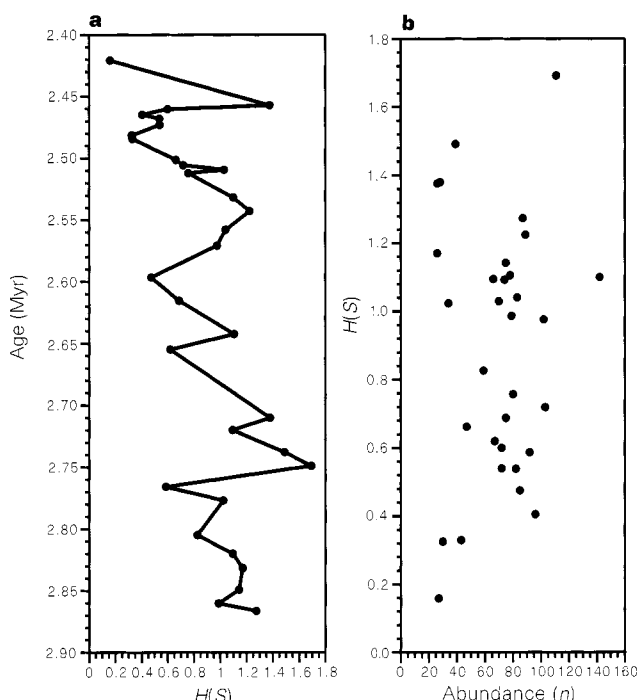


Figure 3 **a**, Deep-sea ostracod species diversity at DSDP site 610. **b**, Relation between ostracod abundance and diversity at site 610.

29. Rex, M. A. *et al. Nature* **365**, 636–639 (1993).

30. Cronin, T. M. *et al. Paleogeography* **10**, 259–281 (1995).

Acknowledgements. We thank the Ocean Drilling Program for samples; H. J. Dowsett and G. Dwyer for advice on site 607 samples; M. A. Buzas, L. C. Hayak and L. Keigwin for discussion about species diversity; and D. Willard, K. Swanson, B. Corliss and M. Rex for comments on the manuscript. Funded by the US Geological Survey Global Change Program (T.M.C.) and a grant from the American Chemical Society (to M.E.R.).

Correspondence and requests for materials should be addressed to T.M.C. (e-mail: tcronin@geochange.er.usgs.gov).

Evolutionary origin of insect wings from ancestral gills

Michalis Averof & Stephen M. Cohen

European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Two hypotheses have been proposed for the origin of insect wings. One holds that wings evolved by modification of limb branches that were already present in multibranching ancestral appendages and probably functioned as gills^{1–5}. The second proposes that wings arose as novel outgrowths of the body wall, not directly related to any pre-existing limbs⁶. If wings derive from dorsal structures of multibranching appendages, we expect that some of their distinctive features will have been built on genetic functions that were already present in the structural progenitors of insect wings, and in homologous structures of other arthropod limbs. We have isolated crustacean homologues of two genes that have wing-specific functions in insects, *pdm* (*nubbin*) and *apterous*. Their expression patterns support the hypothesis that insect wings evolved from gill-like appendages that were already present in the aquatic ancestors of both crustaceans and insects.

Developmental studies have shown that wings and legs originate in a common primordium in early insect embryos⁷ and that they share a common regulatory mechanism for patterning along the antero–posterior (AP) axis⁸, based on the contiguous domain of *engrailed* (*en*) expression at the posterior of each segment. Crustaceans appear to be the closest living relatives of insects that retain a primitively multibranching structure in their limbs^{9,10}, and these limbs show comparable expression of *engrailed*, in a contiguous domain that runs along the posterior of each limb branch (N. Patel and M.A., manuscript in preparation; Fig. 1). Thus, the embryonic origin and relative organization of insect legs and wings appear comparable to those of ventral and dorsal branches of crustacean limbs. Although AP patterning appears to be similar in legs and wings, several genes, including *vestigial*, *scalloped*, *wingless*, *pdm* and *apterous*, have been identified with distinctive roles in establishing the wing primordium and in patterning the wing along the dorso–ventral (DV) axis^{11–17} (Fig. 1).

To determine whether genes with wing-specific functions in insects might play a specific role in patterning dorsal elements of multibranching crustacean appendages, we isolated homologues of the *Drosophila* genes *pdm* and *apterous* (*ap*) from the branchiopod crustacean *Artemia franciscana* (named *Af-pdm* and *Af-ap*, respectively). Amino-acid sequence comparison reveals a high degree of conservation within regions of known sequence motifs and indicates that these are orthologues of the *Drosophila* *pdm* and *apterous* genes (Fig. 2). In *Drosophila* there are two closely related *pdm* genes, *pdm1* and *pdm2*, with largely overlapping expression patterns and functions^{18–20}. We have recovered a single *pdm* gene in *Artemia* (7/7 sequenced clones) with sequence that is approximately equidistant from *pdm1* and *pdm2* (Fig. 2b). In view of the close functional relationship of the two *Drosophila* *pdm* genes^{19,20}, we suspect that *Af-pdm* may be functionally equivalent to both.

One of the *Drosophila* *pdm* genes, *pdm1* (*nubbin*), is expressed throughout the prospective wing and has been implicated in the

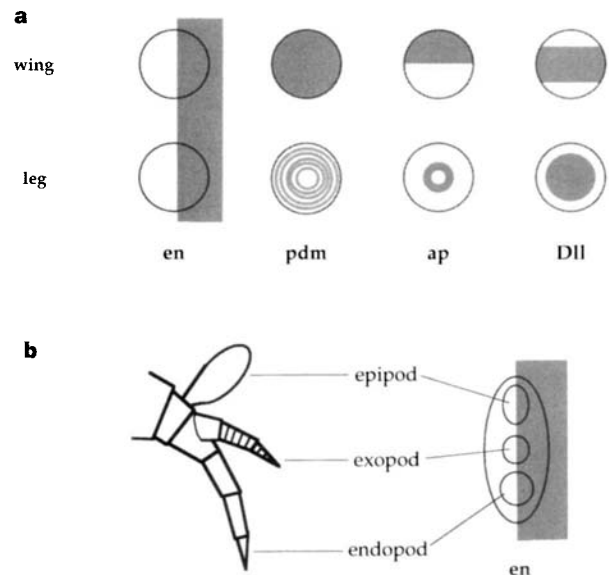


Figure 1 a. Representation of insect legs and wings, indicating the expression of *engrailed* (*en*), *nubbin/pdm*, *apterous* (*ap*), and *Distal-less* (*Dll*) (based on refs 13, 26). The expression of *ap* in the leg appears to have no function²¹. Orientation: dorsal is up, anterior is to the left, and the proximo–distal axis runs perpendicular to the page, with distal regions represented at the centre and proximal regions (attachment to the body wall) at the periphery of each disc. **b.** The multibranching structure of a diagrammatic crustacean limb, indicating its anteroposterior organization relative to the expression of *engrailed* (N. Patel and M.A., manuscript in preparation). The drawing showing *engrailed* expression is oriented as in **a**.

early specification of the wing primordium^{13,14}. It is also expressed more weakly in a set of rings in the primordia of legs, where its function is unknown¹³. To compare the expression of *pdm* in *Artemia* with that seen in *Drosophila*, we have raised antibodies against Af-PDM. Af-PDM is expressed in a dynamic pattern in the developing thoracic limbs (Fig. 3c–e). At early stages, the gene is expressed over most of the developing limb bud. However, as soon as the appendage shows the first signs of regionalization, this expression becomes restricted to a dorsal lobe that will give rise to the distal epipodite. During subsequent stages, expression is maintained specifically in the distal epipodite.

In insects, *apterous* is expressed specifically on the dorsal surface of developing wings^{21,22} and appears to be the primary determinant for DV patterning of wings^{11,16,17}. It is also expressed in a ring in the fourth tarsal segment of *Drosophila* legs, although this expression appears to have no function²¹. Using antibodies raised against Af-AP, we find that this protein is expressed in a pattern similar to that of Af-PDM, with expression in the early limb primordium becoming restricted to the distal epipodite at the time when this structure first becomes morphologically distinct (Fig. 3f–h). Expression is not restricted to the dorsal aspect of this branch; this is not surprising as this lobe shows no sign of dorso–ventral differentiation in its morphology. In addition, at late stages Af-AP expression appears in the limb muscles.

It is of interest that Af-PDM and Af-AP become specifically expressed in the cells of the distal epipodite before these acquire their characteristic differentiated morphology (large nuclei, large intercellular spaces). These expression patterns contrast markedly with that of *Distal-less* (*Dll*), which is expressed in all outgrowing appendages (including insect legs and wings, and all crustacean limb branches^{23,24}; Fig. 3b).

One of our antibody preparations, raised against Af-PDM, fortuitously recognizes a conserved epitope in PDM proteins of

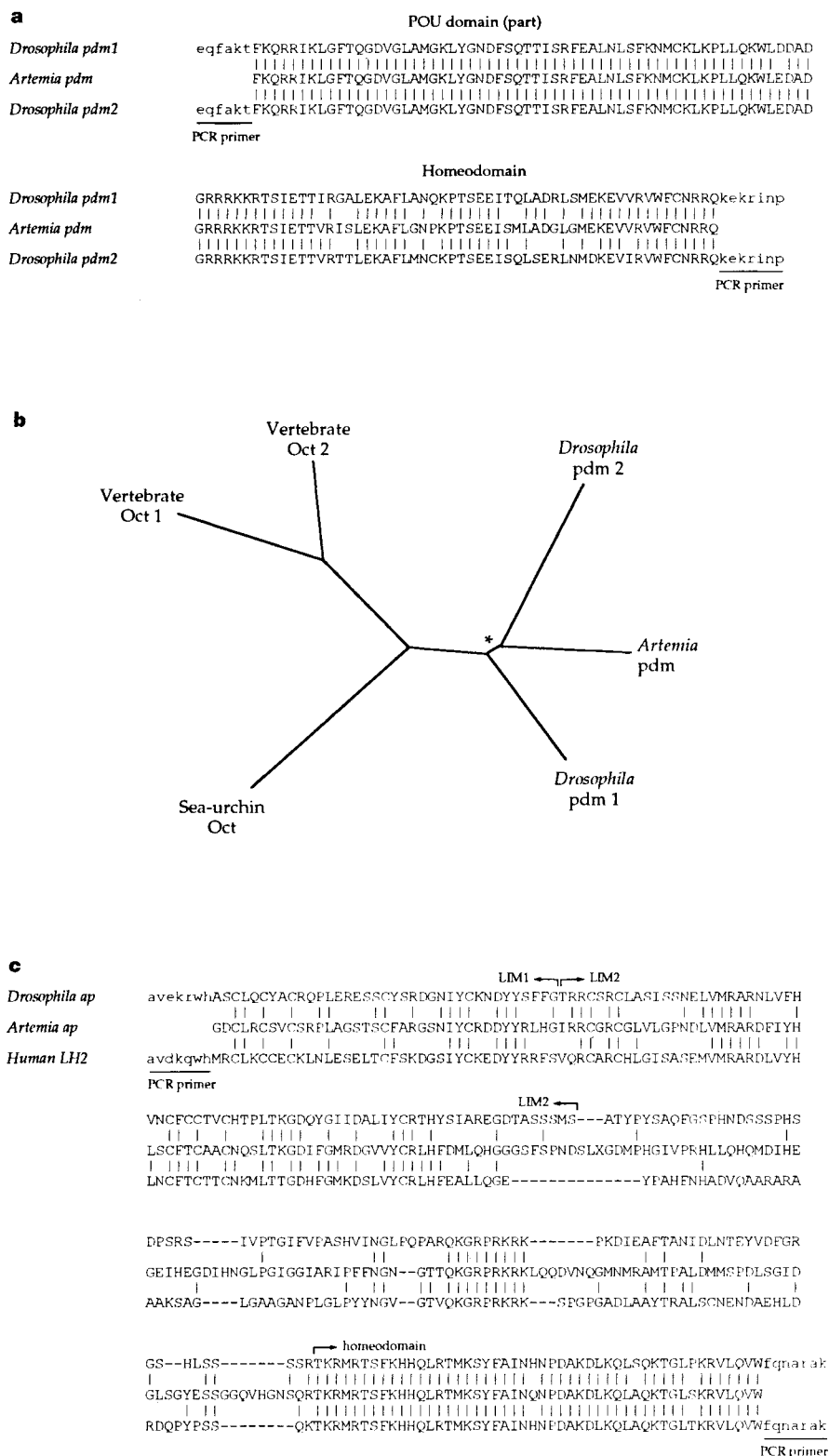


Figure 2 a, Sequence alignment of *Af-PDM* with its *Drosophila* counterparts *Pdm1* and *Pdm2*. Alignments are shown for the POU domain and homeodomain regions only; no significant sequence similarity was found outside these regions. Amino-acid identity is indicated by a vertical line. The positions of PCR primers used for the isolation of the *Artemia* gene (see Methods) are shown. **b**, Tree depicting the sequence relationships of *Af-pdm* with *Drosophila pdm1* and *pdm2*, and their counterparts from vertebrates and sea urchins (Oct genes). Comparisons were made at the amino-acid level on the available POU and homeodomain sequences only (as in **a**). The node between *Af-pdm*, *pdm1* and *pdm2* (marked by

an asterisk) is not significantly resolved. **c**, Sequence alignment of *Af-AP* with its *Drosophila* and vertebrate counterparts *AP* and *LH2*. The entire region available for *Af-AP* is shown, including regions of the two LIM domains (part of *Lim1* and *Lim2*) and part of the homeodomain. Amino-acid identity and the position of PCR primers are indicated as in **a**. Sequence alignments and the tree were prepared using the CLUSTALV program²⁷. Sequence accession numbers: *Af-pdm*, Y09913; *pdm1*, M81957; *pdm2*, M81958; *Oct1*, X13403; *Oct2*, M36653; *Sp-Oct*, L04646; *Af-ap*, Y09914; *ap*, X65158; *LH2*, U11701.

other species (data from *Drosophila*, not shown). We have been able to use this antibody to study *pdm* gene expression in the crayfish *Pacifastacus leniusculus*. In the thoracic limbs of this species, we observe two distinctive patterns of *pdm* expression (Fig. 3i–k): expression throughout a distal epipodite lobe (similar to that seen in *Artemia*) and additional expression in a set of rings along the leg, reminiscent of *pdm* expression in insect legs¹³ (Fig. 1a). *pdm* expression is absent from the endopods of the first thoracic segment in crayfish (Fig. 3i), which do not develop the characteristic articulated morphology of legs, as in *Artemia* limbs (Fig. 3d, e).

We have reasoned that the relationship of insect wings to other structures may be revealed by commonly inherited patterns of gene expression. In this respect, some genes are likely to be more informative than others. Genes like *Dll* and *engrailed* are unlikely

to be informative because they are expressed in similar patterns in all outgrowths from the body wall (refs 24, 25, and N. Patel and M.A., manuscript in preparation), and could have been co-opted into patterning of wings irrespective of their origin. *pdm* and *apterous* appear to be more suitable markers as they have important functions that are more specific to particular appendages. Our observations in *Artemia* indicate that the activities of both *pdm* and *apterous* are associated specifically with a distal epipodite of crustacean limbs. Moreover, we have shown that *pdm* expression is conserved in representatives of two major divergent clades of crustaceans with very different limb architectures, branchiopods and malacostracans. Crustacean epipodites are dorsally located limb branches with respiratory and osmoregulatory functions, precisely the type of structure that, according to some hypotheses, would

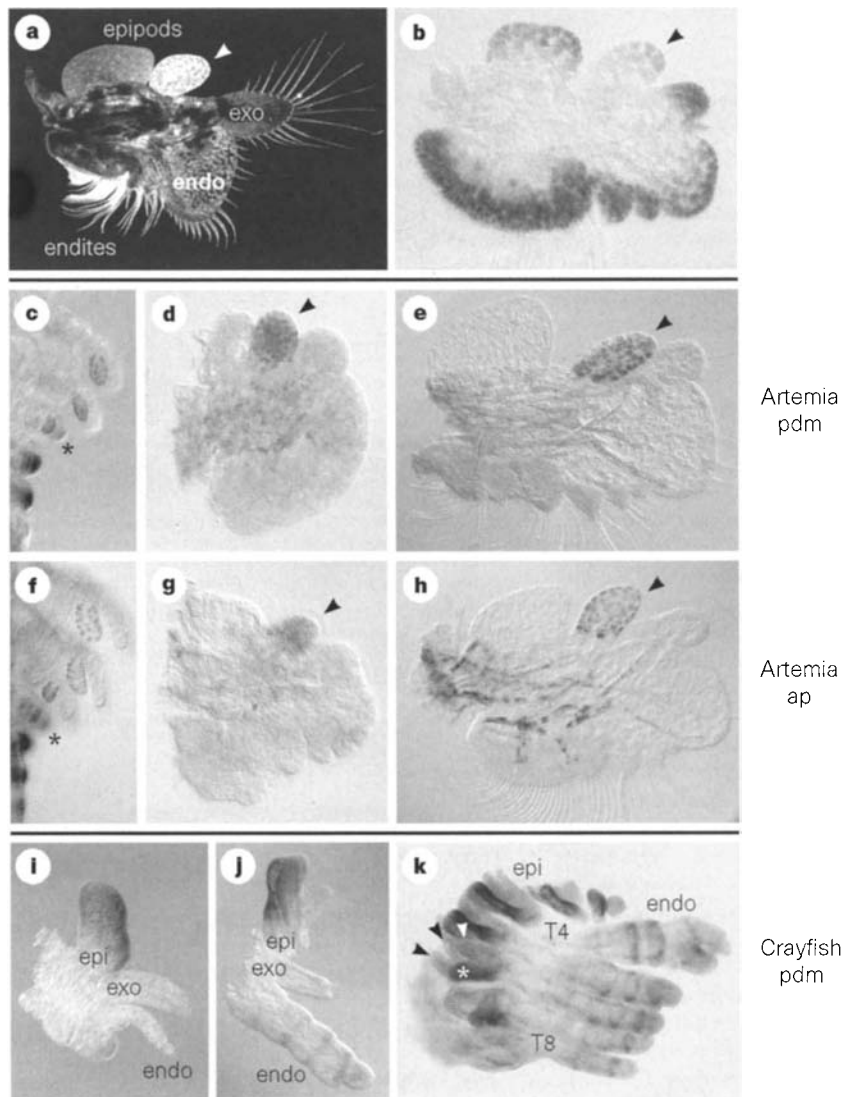


Figure 3 Expression of PDM and AP in developing crustacean appendages. **a**, The branched morphology of an *Artemia* thoracic limb, indicating the position of the two epipodites, the exopod, endopod and endites. Dorsal is up, proximal (attachment to the body wall) is to the left. **b**, Expression of *DLL* in all outgrowing regions of the *Artemia* limb²⁴. **c–e**, Expression of Af-PDM. **f–h**, Expression of Af-AP. **c** and **f**, Dorsal view of a series thoracic appendages in the body of a developing *Artemia* larva. Posterior segments (at the bottom) carry younger limb buds, whereas more anterior appendages (towards the top) represent progressively more mature stages. The stage at which expression becomes restricted to the distal epipodite is marked by an asterisk. **d**, **e** and **g**, **h**, Individual dissected limbs oriented as in **a**. In **d** and **g**, In young limb buds, soon after the earliest regional features (distinct lobes) have become visible. These limbs are of

comparable age to those marked by asterisks in **c** and **f**, **e** and **h**. Later, more mature limbs. Arrowheads mark the distal epipodite. **i–k**, PDM expression in the thoracic limbs of *Pacifastacus leniusculus* at ~70% of embryonic development: **i**, limb of first thoracic segment (T1); **j**, limb of third thoracic segment (T3); **k**, limbs of thoracic segments 4–8 (T4–T8), epipodites of T2 and T3 are also visible anteriorly. Limbs in **i** and **j** are oriented with dorsal side up and proximal towards the left; limbs in **k** are viewed from the dorsal side, oriented with anterior limbs towards the top and proximal parts towards the left. Epipodites, exopods and endopods are labelled where present (exopods absent from T4–T8). Strong staining can be seen in a single, distal and posterior epipodite lobe per limb (lacking in T8) marked by an asterisk in T6 of panel **k**; other epipodite lobes on the same limb are marked by arrowheads.

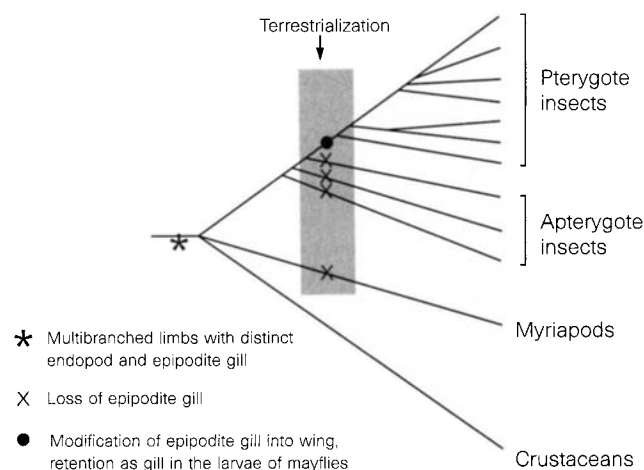


Figure 4 The phylogenetic distribution of respiratory epipodites, wings and gills. Respiratory epipodites are inferred to have been present in the last common ancestor of crustaceans, insects, and possibly myriapods. Thus, the ancestors of insects are likely to have been aquatic animals bearing multibranching appendages with distinct locomotory and respiratory parts (legs and epipodite gills). During the transition from aquatic to terrestrial life epipodite gills must have gradually lost their utility as respiratory and osmoregulatory organs. As a result, gills appear to have been lost independently in myriapods and most of the early lineages of insects (apterygotes). In the lineage that gave rise to the winged insects (pterygotes), we propose that these structures were retained in modified form, perhaps initially as gills in aquatic larvae (as in the abdominal segments of present-day mayfly larvae) and subsequently as wings. The relationships between crustaceans, myriapods and insects remain controversial^{9,10} and are therefore presented as an unresolved trichotomy.

have given rise to the insect wings^{1–5}. We therefore support the idea that insect wings have evolved from gill-like appendages, and propose that they may be homologous to specific epipodites of crustacean limbs.

Homology of divergent structures can never be proven with certainty. Arguments based on gene expression must take into account that individual genes can acquire different roles in different developmental contexts. For example, both *pdm* and *apterous* show distinctive patterns of expression in wings and legs in *Drosophila*^{13,21}. In the case of *pdm*, we have shown that similar distinctive patterns of expression exist in the epipodites and legs (respectively) of crustacean limbs, when these structures exist. Thus we suggest that distinct structural progenitors of legs and epipodites/wings were present in the last common ancestor of crustaceans and insects (Fig. 4). We take this as evidence for a direct evolutionary relationship between epipodites and wings. An alternative interpretation might be that wings do not derive from epipodites but have nevertheless independently coopted a number of gene functions that were already used in epipodites. Although formally possible, we consider this less likely.

The proposed evolution of wings from epipodites presumably involved several genetic and developmental changes, required for repatterning of these structures. At the genetic level, these changes must have utilized and modified gene functions that were already operating in respiratory epipodites. One example may be the use of *apterous* expression and its restriction to the dorsal surface of the wing to initiate a DV patterning system not used in the ancestral limb or in present-day insect legs^{16,17}. A second modification must have resulted in the physical separation of the epipodite from the leg, perhaps by fusion of the limb base with the body wall^{3,5}. □

Methods

Af-pdm and *Af-ap* were cloned by polymerase chain reaction (PCR) using degenerate primers. Primers were 5'-GGAATTC GA(A/G) CA(A/G) TT(T/C)

GCI AA(A/G) AC-3' and 5'-GCTCTAGA GG(A/G)TT IAT IC(T/G) (T/C)TT (T/C)TC (T/C)TT-3' for *Af-pdm*, and 5'-GGAATTC GCI GTI GAI AA(A/G) C(A/G)I TGG CA-3' and 5'-GCTCTAGA TT IGC IC(T/G) IGC (A/G)TT(T/C)TG (A/G)AA-3' for *Af-ap* (where I is inosine). PCRs were carried out on first-strand cDNA prepared from early larval stages of *Artemia franciscana*; 7 and 4 independent clones were sequenced for *Af-pdm* and *Af-ap* respectively. *Af-pdm* and *Af-ap* were subcloned into expression vectors pATH-1 and pET-23a and expressed in bacteria. Antibodies against AF-PDM and Af-AP were raised in mice by repeated injection of 70-μg doses of bacterially expressed protein in Freund's adjuvant. Immunochemical staining was done using 1:500 dilutions of mouse serum. *Artemia* larvae were sonicated before staining to allow efficient penetration of reagents. The DLL staining in Fig. 3b was done with the antibody described in ref. 24. Detailed protocols are available upon request.

Received 31 October; accepted 12 December 1996.

- Wigglesworth, V. B. *Nature* **246**, 127–129 (1973).
- Wigglesworth, V. B. in *Insect Flight* (ed. Rainey, R. C.) 255–269 (Blackwell Scientific, Oxford, 1976).
- Kukalova-Peck, J. *Can. J. Zool.* **61**, 1618–1669 (1983).
- Kukalova-Peck, J. in *The Insects of Australia* (ed. Naumann, I. D.) 141–179 (Melbourne University Press, CSIRO, 1991).
- Kukalova-Peck, J. *Can. J. Zool.* **70**, 236–255 (1992).
- Snodgrass, R. E. *Principles of Insect Morphology* (McGraw Hill, New York, 1935).
- Cohen, B., Simcox, A. A. & Cohen, S. M. *Development* **117**, 597–608 (1993).
- Basler, K. & Struhl, G. *Nature* **368**, 208–214 (1994).
- Averof, M. & Akam, M. *Phil. Trans. R. Soc. Lond. B* **347**, 293–303 (1995).
- Friedrich, M. & Tautz, D. *Nature* **376**, 165–167 (1995).
- Williams, J. A., Paddock, S. W. & Carroll, S. B. *Development* **117**, 571–584 (1993).
- Williams, J. A., Paddock, S. W., Vorwerk, K. & Carroll, S. B. *Nature* **368**, 299–305 (1994).
- Ng, M., Diaz-Benjumea, F. J. & Cohen, S. M. *Development* **121**, 589–599 (1995).
- Ng, M., Diaz-Benjumea, F. J., Vincent, J.-P., Wu, J. & Cohen, S. M. *Nature* **381**, 316–319 (1996).
- Kim, J. *et al.* *Nature* **382**, 133–138 (1996).
- Diaz-Benjumea, F. J. & Cohen, S. M. *Cell* **75**, 741–752 (1993).
- Blair, S. S., Brower, D. L., Thomas, J. B. & Zavortink, M. *Development* **120**, 1805–1815 (1994).
- Lloyd, A. & Saconju, S. *Mech. Dev.* **36**, 87–102 (1991).
- Bhat, K. M., Poole, S. J. & Schedl, P. *Mol. Cell. Biol.* **15**, 4052–4063 (1995).
- Yeo, S. L. *et al.* *Genes Dev.* **9**, 1223–1236 (1995).
- Cohen, B., McGuffin, M. E., Pfeifle, C., Segal, D. & Cohen, S. M. *Genes Dev.* **6**, 715–729 (1992).
- Carroll, S. B. *et al.* *Science* **265**, 109–114 (1994).
- Cohen, S. M., Brönnner, G., Küttner, E., Jürgens, G. & Jäckle, H. *Nature* **338**, 432–434 (1989).
- Panganiban, G., Sebring, A., Nagy, L. & Carroll, S. B. *Science* **270**, 1363–1366 (1995).
- Patel, N. H., Kornberg, T. B. & Goodman, C. S. *Development* **107**, 201–212 (1989).
- Cohen, S. M. in *Drosophila Development* (eds Martinez-Arias, A. & Bate, M.) 747–841 (Cold Spring Harbor Press, Cold Spring Harbor, New York, 1993).
- Higgins, D. G., Bleasby, A. J. & Fuchs, R. *CABIOS* **8**, 189–191 (1991).

Acknowledgements. We thank D. Holdich for crayfish embryos, G. Panganiban for antibody against DLL, N. Patel for communicating unpublished observations, and M. Akam, F. Ferrari, T. Lecuit and D. Walossek for comments and discussion. M.A. is supported by the Human Capital and Mobility Programme of the European Community.

Correspondence and requests for materials should be addressed to M.A. (averof@embl-heidelberg.de). The *Af-pdm* and *Af-ap* sequences have been deposited in the EMBL database under accession numbers Y09913 and Y09914.

Use-dependent increases in glutamate concentration activate presynaptic metabotropic glutamate receptors

Massimo Scanziani^{†*}, Paul A. Salin^{†*}, Kaspar E. Vogt^{*}, Robert C. Malenka^{‡§} & Roger A. Nicoll^{†*}

Departments of ^{*} Cellular and Molecular Pharmacology, [‡] Physiology, and [§] Psychiatry, University of California at San Francisco, San Francisco, California 94143, USA

The classical view of fast chemical synaptic transmission is that released neurotransmitter acts locally on postsynaptic receptors and is cleared from the synaptic cleft within a few milliseconds by

[†] Present addresses: Laboratoire de Physiologie, Ecole Supérieure de Physique et Chimie Industrielles, 10 rue Vauquelin, 75005 Paris, (M. S.); Institut Alfred Fessard, Av. de la Terrasse, 91198 Gif sur Yvette cedex, France (P. A. S.).

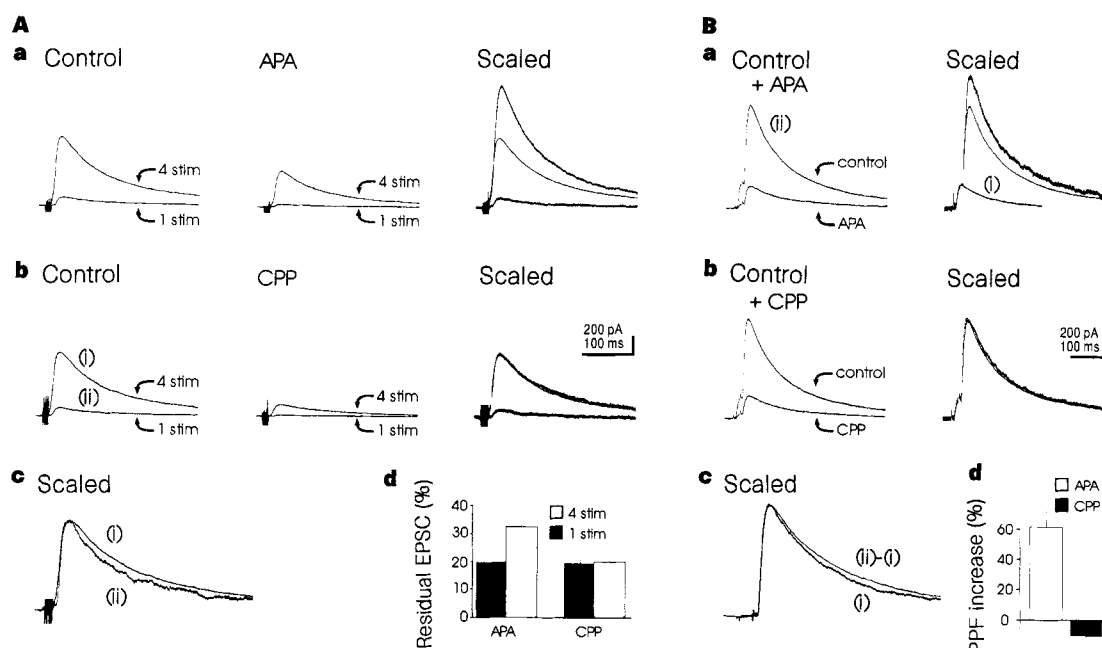


Figure 1 Reduced antagonism of mossy fibre NMDAR EPSCs by the competitive low-affinity NMDAR antagonist APA during brief high-frequency stimulation. **A, a**, A mossy fibre input was stimulated at 0.05 Hz by alternating a single stimulus with four stimuli at 3-ms intervals (superimposed sweeps). The amplitude of the EPSC evoked with 4 stimuli averaged $530 \pm 93\%$ of the amplitude of the EPSC evoked with a single stimulus ($n = 6$). Application of APA (1 mM) reduced the EPSC evoked by a single stimulus to a greater extent than the EPSC evoked by four stimuli. **b**, Same neuron as in **a**. The high-affinity NMDAR antagonist CPP (1 μ M) reduced both EPSCs to the same extent. **c**, The decay of the EPSC evoked by a single stimulus was faster than the decay of the EPSC evoked by four stimuli. The peak of the EPSC evoked by a single stimulus (see **b**, control) was scaled to the peak of the EPSC evoked by four stimuli. The decay of the EPSC increased from 62 ± 3 ms to 79 ± 5 ms ($n = 6$). **d**, Summary graph of the amplitude of the EPSC recorded in the presence of a non-saturating concentration of the competitive

antagonist APA (1 mM; $n = 6$) and CPP (1 μ M; $n = 3$) (residual EPSC). **B, a**, A mossy fibre input was stimulated at 0.05 Hz by alternating a single stimulus (i) with two stimuli separated by 20 ms (ii). The paired-pulse facilitation (PPF) in control conditions averaged $262 \pm 79\%$ ($n = 6$) (see Methods for calculations). APA (1 mM) reduced the first EPSC to a greater extent than the second, increasing PPF, as shown by scaling of the amplitudes of the first EPSCs in the absence and presence of APA. **b**, Same neuron as in **a**. CPP (700 nM) reduced both first and second EPSCs to a similar extent. **c**, Decay of the second EPSC (trace (ii) in **a**) after subtraction of the single EPSC (trace (i) in **a**) was slower than the decay of the single EPSC. The decay of the EPSC increased from 60 ± 4 ms to 68 ± 3 ms ($n = 6$). **d**, The change in the ratio (PPF) between the amplitude of the second EPSC (estimated after subtraction of the single EPSC) and the amplitude of the single EPSC is plotted for APA (1 mM, $n = 6$) and CPP (0.7–1 μ M; $n = 4$).

diffusion and by specific reuptake mechanisms. This rapid clearance restricts the spread of neurotransmitter and, combined with the low affinities of many ionotropic receptors, ensures that synaptic transmission occurs in a point-to-point fashion¹. We now show, however, that when transmitter release is enhanced at hippocampal mossy fibre synapses, the concentration of glutamate increases and its clearance is delayed; this allows it to spread away from the synapse and to activate presynaptic inhibitory metabotropic glutamate receptors (mGluRs). At normal levels of glutamate release during low-frequency activity, these presynaptic receptors are not activated. When glutamate concentration is increased by higher-frequency activity or by blocking glutamate uptake, however, these receptors become activated, leading to a rapid inhibition of transmitter release. This effect may be related to the long-term depression of mossy fibre synaptic responses that has recently been shown after prolonged activation of presynaptic mGluRs (refs 2, 3). The use-dependent activation of presynaptic mGluRs that we describe here thus represents a negative feedback mechanism for controlling the strength of synaptic transmission.

Previous studies of cochlear synapses in the chick brain stem⁴, cerebellar synapses⁵ and cultured neurons⁶ indicate that when release is enhanced, the concentration of glutamate can increase and spread from the synapse. If a similar process were to occur at cortical synapses⁷, it could have important consequences on the input specificity of these synapses, as well as on their use-dependent modification. We tested for possible use-dependent increases in glutamate concentration at mossy fibre synapses, which contact the

proximal dendrites of CA3 pyramidal cells, by using the low-affinity, rapidly dissociating, competitive N-methyl-D-aspartate receptor (NMDAR) antagonist amino pimelic acid (APA)^{8,9} and comparing its ability to block NMDAR-mediated excitatory postsynaptic currents (NMDAR EPSCs) evoked by single versus repetitive stimuli, which greatly facilitate release of glutamate¹⁰. With such an antagonist, the degree of reduction of NMDAR EPSCs will decrease as the concentration and/or time course of glutamate in the synaptic cleft increase because the higher concentration of glutamate can rapidly displace the low-affinity antagonist from the receptors⁶.

In these experiments, AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate) receptors were blocked with NBQX (6-nitro-7-sulphamoylbenzo[*f*]quinoxaline-2,3-dione) (10 μ M), the cells were held at +30 to +40 mV to record the NMDAR EPSCs, and the ability of APA to antagonize the synaptic response to one versus four stimuli (with an interstimulus interval of 3 ms) was compared (Fig. 1A, a). When the size of the single-stimulus response in APA is scaled up and superimposed on the control single stimulus response (scaled), it is clear that the response to four stimuli is considerably larger than the control response, indicating that APA was less effective at blocking the facilitated response. On the other hand, if in the same cell the high-affinity antagonist CPP (D-carboxypiperazin-propyl-phosphonic acid) is applied, the degree of antagonism is identical for both responses, as shown by the superimposition of the scaled responses (Fig. 1A, b and d). Similar results were obtained when only two rather than four stimuli were delivered (Fig. 1B). We

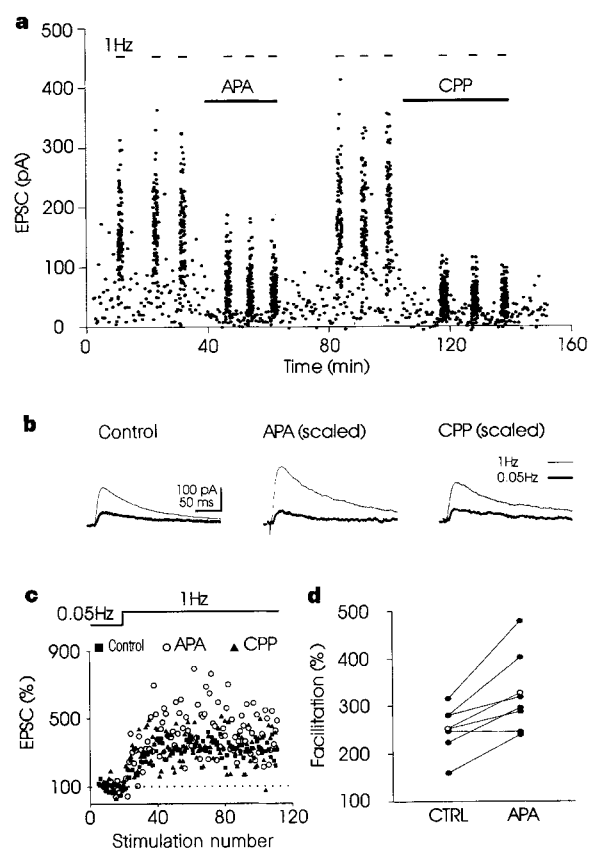


Figure 2 Decreased antagonism of facilitated NMDAR EPSCs by APA. **a**, Mossy fibres were stimulated alternately at 0.05 and 1 Hz. Application of APA (500 μ M) inhibited the EPSC evoked at 0.05 Hz to a greater extent than the EPSC evoked at 1 Hz. CPP (250 nM) inhibited EPSCs evoked at 0.05 Hz and at 1 Hz to a similar extent. **b**, Data from the experiment in **a**. The responses at 0.05 Hz in APA or CPP have been scaled to the control responses at 0.05 Hz. **c**, Time course of facilitation of the experiment in **a** after normalization of EPSC amplitudes evoked at 0.05 Hz under control conditions or in the presence of either APA or CPP. Note the enhanced facilitation in the presence of APA. **d**, Scatter plot of 7 experiments. Open circles represent average values.

also found that decay of the NMDAR EPSC is slowed when transmitter release is increased by repetitive stimulation (Fig. 1A, c and B, c), consistent with the delayed clearance of glutamate from the synaptic region. Four stimuli prolonged decay by $29 \pm 9\%$ ($n = 6$) ($P < 0.02$) and two stimuli prolonged it by $14 \pm 3\%$ ($n = 6$) ($P < 0.01$).

Although our results with APA are consistent with an increase in glutamate concentration during repetitive stimuli, a limitation of these experiments is that under control conditions the increase in the amplitude of the EPSC due to two or more stimuli might be limited by receptor occupancy, if glutamate release occurs at the same site during repeated stimulation. In the presence of APA, a fraction of receptors will be protected during the first stimulus, and because of the fast re-equilibration of APA, be available to subsequent stimuli, leading to the observed results. To circumvent this problem, we have used frequency facilitation to enhance transmitter release. As mossy fibres show markedly enhanced release at frequencies at which the NMDAR EPSC completely recovers before the next stimulus, the limitation described above is avoided. As shown in Fig. 2a, mossy fibres were stimulated at 0.05 Hz and then given bouts of 1-Hz stimulation in control conditions and following brief applications of APA or CPP. When the responses are normalized to those recorded during low-frequency stimulation (0.05 Hz) (Fig. 2b, c), it is evident that the 1-Hz responses in CPP scale to the control responses, but that the 1-Hz responses are less inhibited by

APA. This decreased effectiveness of APA on 1-Hz responses was seen in six of seven experiments (Fig. 2d). APA significantly increased facilitation from $251 \pm 19\%$ to $324 \pm 33\%$ ($P < 0.02$) ($n = 6$), whereas CPP had no significant effect ($P > 0.4$) ($n = 4$).

Having established that the concentration of glutamate rises in the synaptic region during repetitive stimulation, we next examined whether this might result in the activation of presynaptic metabotropic glutamate receptors (mGluRs). mGluRs are present on a large number of presynaptic terminals in the mammalian central nervous system (CNS)^{11–16}, including mossy fibre terminals^{2,3,17–21}, and when activated by exogenously applied agonists cause a depression of transmitter release. However, the conditions necessary for synaptically released glutamate to gain access to these presynaptic receptors are unknown. In Fig. 3A, a, mossy fibres were stimulated at a frequency of 0.05 Hz and bouts of 1-Hz stimulation were applied. Application of the broad-spectrum mGluR antagonist MCPG ((+)- α -methyl-4-carboxyphenylglycine), which is known to antagonize presynaptic mGluRs on mossy fibre synapses¹⁹, had no effect on the responses evoked at low frequency. However, the responses evoked at 1 Hz were enhanced, suggesting that during the 1-Hz stimulation glutamate was able to activate the presynaptic inhibitory receptors. A summary of the effect of MCPG in five similar experiments is shown in Fig. 3A, b and c. A similar enhancement was observed in seven of seven experiments with MCCG ((2S,3S,4S)-methyl-2-(carboxycyclopropyl)-glycine) (500 μ M), which selectively antagonizes the presynaptic action of group-2 mGluR agonists at mossy fibre synapses¹⁹.

It could be argued that mGluRs are activated at low-frequency stimulation but that the duration of the presynaptic inhibition is shorter than the interstimulus interval. Alternatively, the mGluRs might require a certain level of repetitive synaptic activation to inhibit transmission. This latter possibility can be ruled out because the inhibition induced by bath-applied mGluR agonists (not shown) was independent of the stimulus frequency. To test the former possibility, we examined the effects of MCPG when stimuli were applied at shorter interstimulus intervals. We chose an interval of 200 ms as this is the optimal interval for presynaptic inhibition mediated by GABA_B receptors²². MCPG had no consistent effect on the second response to paired-pulse stimulation (Fig. 3B, a), suggesting that the amount of glutamate released following a single stimulus is unable to activate the mGluRs. If, however, in this same experiment, an additional stimulus was applied immediately before the first, which greatly increased transmitter release, the test response was now clearly enhanced by MCPG (Fig. 3B, a; right traces). A summary graph of five similar experiments (Fig. 3B, b and c) confirms that single stimuli at low-release probability fail to activate mGluRs, but that when release probability is enhanced by paired pulse, stimulation mGluRs are activated.

The need for increased glutamate release for mGluR activation could mean that mGluRs are located within the synaptic cleft and require an increase in glutamate concentration, or that mGluRs are not present in the cleft and require the spread of glutamate from the release site for their activation. Given that the affinity of glutamate for group-2 mGluRs, $\sim 10 \mu$ M (ref. 12), is higher than for AMPA receptors, typically greater than 100 μ M (ref. 23), the first situation is unlikely. If mGluR activation depends on spread of glutamate from the release site, the activation of these receptors should be enhanced by blocking the uptake of glutamate. Indeed, when the glutamate-uptake blocker L-trans-pyrrolidine-2,4-dicarboxylic acid (trans-PDC)²⁴ is applied, the size of the field EPSPs evoked at 1-Hz stimulation, but not at 0.017-Hz stimulation, is reduced (Fig. 4a). This depression is entirely reversed by MCPG, and, in fact, is replaced by an enhancement over the control responses evoked at 1-Hz stimulation (Fig. 4b). Results were similar in four other experiments (Fig. 4c), strongly supporting a model in which glutamate spread from the release site is required for mGluR activation.

In summary, we have found that the concentration of glutamate

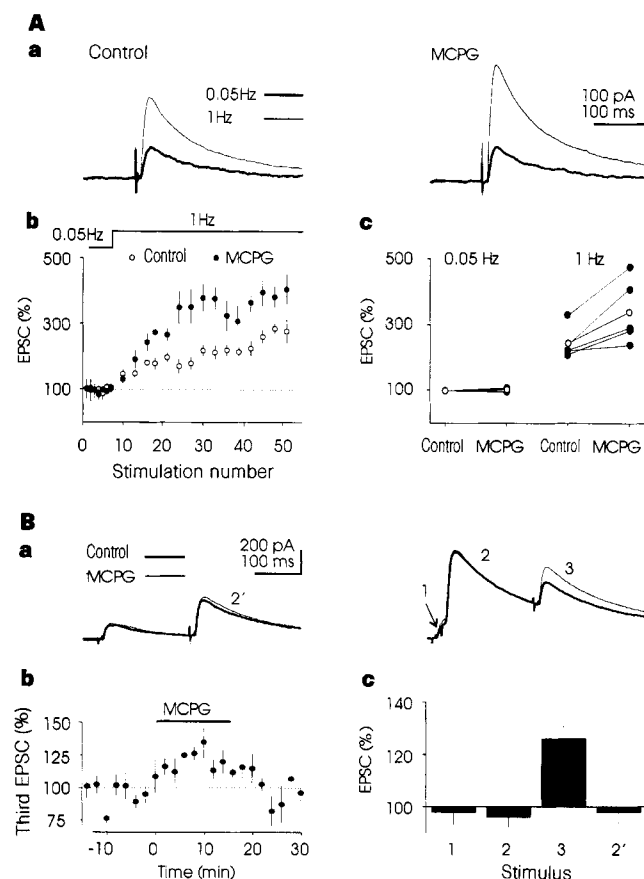
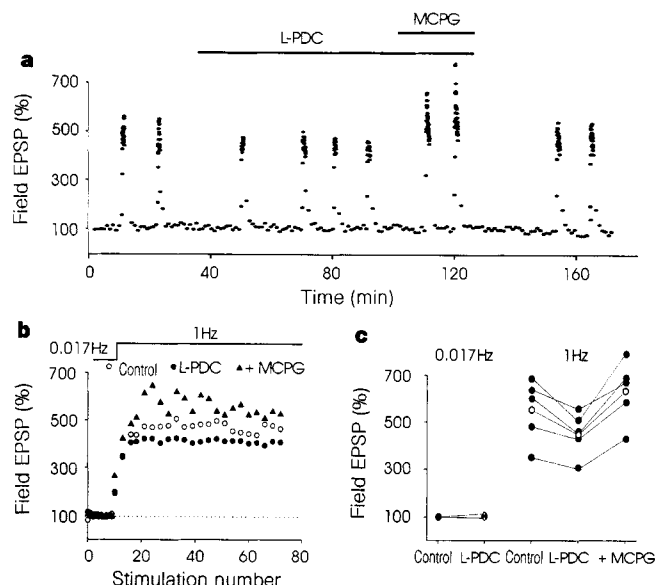


Figure 3 Frequency-dependent activation of presynaptic metabotropic glutamate receptors. **A, a**, Mossy fibres were stimulated alternately at 0.05 and 1 Hz (superimposed sweeps). Perfusion of the metabotropic glutamate receptor antagonist MCPG (1.5 mM) had no effect on baseline transmission (0.05 Hz) but increased the mean amplitude of the facilitated response evoked at 1 Hz. **b**, Summary of the time course of facilitation for 5 experiments normalized over the 0.05-Hz stimulation period. **c**, Scatter plot of the 5 experiments. MCPG significantly increased the facilitated response from $247 \pm 22\%$ to $343 \pm 50\%$ ($P < 0.03$). **B, a**, Mossy fibres were stimulated at 0.033 Hz with paired stimuli separated by 200 ms, and for every other stimulus cycle, a third stimulus preceded the first by 20 ms (response labelled 1 in the traces on the right). MCPG (1.5 mM) had no effect when paired stimuli were given (response labelled 2' in left-hand traces), but increased the amplitude of the EPSC when preceded by a double stimulus (response labelled 3 in right-hand traces). **b**, Summary of the time course of the action of MCPG on the third EPSC. **c**, Summary of the mean amplitude changes for all EPSCs in the presence of MCPG (0.5–1.5 mM; $n = 8$ for 1, 2 and 3; $n = 3$ for 2'). MCPG increased 3 by $26 \pm 6\%$. In **A, b**, the bin size of each data point during 1-Hz stimulation is 3 stimuli. In **A, c**, open circles represent average values.

in the mossy fibre synaptic cleft is critically dependent on the frequency of synaptic activation. The mossy fibre synapses are ideally suited for frequency-dependent changes in glutamate concentration because they show dramatic frequency facilitation¹⁰ and a single bouton forms multiple release sites^{25,26}, an anatomical arrangement favouring spillover onto adjacent synapses (see also refs 4, 5). A consequence of the spread of glutamate from the release site during certain patterns of afferent activity is that mGluRs, which normally are silent during low-frequency mossy fibre transmission, become activated and exert a negative feedback on trans-

mitter release. This proposal is supported by anatomical evidence indicating that group-2 mGluRs, which we have shown contribute to presynaptic inhibition, are localized at the preterminal zone of mossy fibre synapses away from the release sites^{2,27}. The use-dependent engagement of presynaptic mGluRs provides a local means for controlling the level of release at which mossy fibre synapses can operate. Furthermore, our findings emphasize that receptor localization relative to transmitter release sites is important and influences the ways in which receptors are engaged during synaptic activity. □

Figure 4 Activation of presynaptic metabotropic glutamate receptor by inhibiting glutamate uptake. **a**, Field recordings of mossy fibre EPSPs evoked alternately at stimulation frequencies of 0.017 and 1 Hz. The glutamate-uptake blocker L-PDC (30 μ M) had no effect on the amplitude of field EPSPs evoked at 0.017 Hz but reduced the average amplitude of the field EPSPs evoked at 1 Hz. This effect was reversed by perfusion of MCPG. **b**, Data from experiment in **a**: time course of facilitation under control conditions or in the presence of L-PDC and MCPG. Note that in the presence of MCPG, facilitation overshoots control values. **c**, Scatter plot of 5 experiments. Average values are 554 ± 61 , 448 ± 43 and $634 \pm 60\%$ for the facilitated responses in control, L-PDC, and L-PDC with MCPG, respectively. The differences between control and L-PDC, L-PDC and L-PDC with MCPG, and control and L-PDC with MCPG are significant ($P < 0.02$; $P < 0.01$; $P < 0.005$). In **b**, the bin size of each data point during 1-Hz stimulation is 3 stimuli. In **c**, open circles represent average values.



Methods

Standard procedures were used to prepare hippocampal slices (400–500 μm) from 5–15-day-old guinea-pigs. After dissection, slices were maintained at room temperature for at least 1 h in a submerged chamber containing artificial cerebrospinal fluid (ACSF) equilibrated with 95% O_2 and 5% CO_2 , and then transferred to a superfusing chamber. The ACSF contained (in mM) 119 NaCl, 2.5 KCl, 1 NaH_2PO_4 , 1.3 MgSO_4 , 2.5 CaCl_2 , 26 NaHCO_3 , 11 glucose. For electrophysiological recording, the concentration of CaCl_2 and MgSO_4 was raised to 4 mM. Whole-cell recordings were made at room temperature in the presence of picrotoxin (100 μM) and NBQX (10 μM). Whole-cell recording electrodes were filled with a solution containing (in mM) 122.5 caesium gluconate, 10 CsCl, 10 HEPES, 10 BAPTA, 8 NaCl, 2 Mg-ATP, 0.3 Na-GTP. The resistance of the patch pipette ranged between 2 and 4 M Ω . Access resistances ranged between 4 and 10 M Ω and were not allowed to vary by more than 15% during the course of the experiment. No series resistance compensation was used. In 5 cells we tested the effect of APA (0.5 mM) on AMPAR EPSCs. Picrotoxin was added to isolate the AMPAR currents and a low concentration of CNQX (0.5 μM) was used to prevent epileptiform activity. APA (0.5 mM) had little effect on AMPAR EPSCs ($90 \pm 5\%$ of control; $n = 5$) or on paired-pulse facilitation (PPF). The PPF was 2.7 ± 0.4 before APA and 2.9 ± 0.5 during APA. Field recordings were made with a glass electrode placed in the stratum lucidum in the absence of picrotoxin and NBQX. Bipolar steel electrodes were placed in the granule cell layer to stimulate mossy fibres. Only EPSCs depressed by more than 60% or field EPSCs depressed by more than 80% during application of the mGluR agonist L-AP4 (10 μM), which selectively blocks mossy fibre responses in the guinea-pig^{17–19}, were considered as mossy fibre responses²⁸. This was tested in all experiments. Data were acquired and analysed as described previously²⁸. Average values are expressed as mean $\pm$ s.e.m. Significance was determined by a two-tailed paired t -test, unless stated otherwise. PPF was defined as $P_2 - P_1 / P_1 \times 100$, where P_1 and P_2 are the peak amplitudes of the first (i) and second (ii) EPSC, respectively. P_2 was estimated after subtraction of (i) from (ii) and P_1 was set as the peak amplitude of (i). To determine changes in the decay of the NMDAR EPSC, the first 100 ms of the EPSC was fitted with a single exponential. Drugs used were: caesium hydroxide, D-gluconic acid (Aldrich), picrotoxin, BAPTA, L-AP4 (Sigma), CPP, NBQX, (+)-MCPG (Tocris Cookson), DL- α -amino pimelic acid (ICN Biomedicals), L-trans-pyrrolidine-2,4-dicarboxylic acid (Research Biochemicals International).

Received 15 October; accepted 14 December 1996.

- Hille, B. *Neuron* **9**, 187–195 (1992).
- Yokoi, M. *et al. Science* **273**, 645–647 (1996).
- Kobayashi, K., Manabe, T. & Takahashi, T. *Science* **273**, 648–650 (1996).
- Otis, T. S., Wu, Y.-C. & Trussell, L. O. *J. Neurosci.* **16**, 1634–1644 (1996).
- Silver, R. A., Cull-Candy, S. G. & Takahashi, T. *J. Physiol.* **494**, 231–250 (1996).
- Tong, G. & Jahr, C. E. *Neuron* **12**, 51–59 (1994).
- Kullmann, D. M., Erdemli, G. & Asztely, F. *Neuron* **17**, 461–474 (1996).
- Olverman, H. J., Jones, A. W., Mewett, K. N. & Watkins, J. C. *Neuroscience* **26**, 17–31 (1988).
- Bergles, D. E. & Jahr, C. E. *Neurosci. Abstr.* **22**, 797 (1996).
- Salin, P. A., Scanziani, M., Malenka, R. C. & Nicoll, R. A. *Proc. Natl Acad. Sci. USA* **93**, 13304–13309 (1996).
- Nakanishi, S. *Neuron* **13**, 1031–1037 (1994).
- Pin, J.-P. & Duvosin, R. *Neuropharmacology* **34**, 1–26 (1995).
- Conn, P. J., Winder, D. G. & Gereau, R. W. I. In *The Metabotropic Glutamate Receptors* (eds Conn, P. J. & Patel, J.) (Humana, Totowa, NJ, 1994).
- Forsythe, I. D. & Clements, J. D. *J. Physiol.* **429**, 1–16 (1990).
- Baskys, A. & Malenka, R. C. *J. Physiol.* **444**, 687–701 (1991).
- Glaum, S. R. & Miller, R. J. *J. Neurophysiol.* **70**, 2669–2672 (1993).
- Yamamoto, C., Sawada, S. & Takada, S. *Exp. Brain Res.* **51**, 128–134 (1983).
- Lanthorn, T. H., Ganong, A. H. & Cotman, C. W. *Brain Res.* **290**, 174–178 (1984).
- Manzoni, O. J., Castillo, P. E. & Nicoll, R. A. *Neuropharmacology* **34**, 965–971 (1995).
- Hsia, A. Y. *et al. Neuropharmacology* **34**, 1567–1572 (1995).
- Kamiya, H., Shinozaki, H. & Yamamoto, C. *J. Physiol.* **493**, 447–455 (1996).
- Isaacson, J. S., Solis, J. M. & Nicoll, R. A. *Neuron* **10**, 165–175 (1993).
- Patneau, D. K. & Mayer, M. L. *J. Neurosci.* **10**, 2385–2399 (1990).
- Bridges, R. J., Stanley, M. S., Anderson, M. W., Cotman, C. W. & Chamberlin, A. R. *J. Med. Chem.* **34**, 717–725 (1991).
- Hamlyn, L. H. *J. Anat.* **96**, 112–126 (1962).
- Chicurel, M. E. & Harris, K. M. *J. Comp. Neurol.* **323**, 1–14 (1992).
- Shigemoto, R. *et al. Soc. Neurosci. Abstr.* **21**, 846 (1995).
- Weisskopf, M. G. & Nicoll, R. A. *Nature* **376**, 256–259 (1995).

Acknowledgements. We thank M. Frerking for comments on the manuscript, and H. Czerwonka for secretarial assistance. M.S. is supported by the CIBA-GEIGY-JUBILAEUMS-STIFTUNG. R.A.N. is member of the Keck Center for Integrative Neuroscience and the Silvio Conte Center for Neuroscience Research. R. C. M. is a member of the Center for Neurobiology and Psychiatry. R.A.N. and R.C.M. are supported by grants from the NIH.

Correspondence and requests for materials should be addressed to R.A.N. (e-mail: nicoll@phy.ucsf.edu).

Decreased prefrontal dopamine D1 receptors in schizophrenia revealed by PET

Yoshiro Okubo^{††}, Tetsuya Suhara[†], Kazutoshi Suzuki[†], Kaoru Kobayashi[†], Osamu Inoue[†], Omi Terasaki^{††}, Yasuhiro Someya^{††}, Takeshi Sassa^{††}, Yasuhiko Sudo[†], Eisuke Matsushima[†], Masaomi Iyo[†], Yukio Tateno[†] & Michio Toru^{*}

^{*} Department of Neuropsychiatry, Tokyo Medical and Dental University School of Medicine, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113, Japan

[†] Division of Advanced Technology for Medical Imaging, National Institute of Radiological Sciences, 4-9-1, Anagawa, Inage-ku, Chiba 263, Japan

Schizophrenia is believed to involve altered activation of dopamine receptors, and support for this hypothesis comes from the antipsychotic effect of antagonists of the dopamine D2 receptor (D2R)¹. D2R is expressed most highly in the striatum, but most of the recent positron emission tomography (PET) studies have failed to show any change in D2R densities in the striatum of schizophrenics^{2–5}, raising the possibility that other receptors may also be involved. In particular, the dopamine D1 receptor (D1R), which is highly expressed in the prefrontal cortex⁶, has been implicated in the control of working memory^{7,8}, and working memory dysfunction is a prominent feature of schizophrenia⁹. We have therefore used PET to examine the distribution of D1R and D2R in brains of drug-naïve or drug-free schizophrenic patients. Although no differences were observed in the striatum relative to control subjects, binding of radioligand to D1R was reduced in the prefrontal cortex of schizophrenics. This reduction was related to the severity of the negative symptoms (for instance, emotional withdrawal) and to poor performance in the Wisconsin Card Sorting Test¹⁰. We propose that dysfunction of D1R signalling in the prefrontal cortex may contribute to the negative symptoms and cognitive deficits seen in schizophrenia.

We examined D1R and D2R in 17 male schizophrenics and 18 healthy male volunteers. Ten of the patients were completely naïve for neuroleptics and seven were drug-free for a minimum of two weeks before PET examination. Two PET runs were done in each subject on the same day using [¹¹C]SCH23390 and [¹¹C]N-methyl-

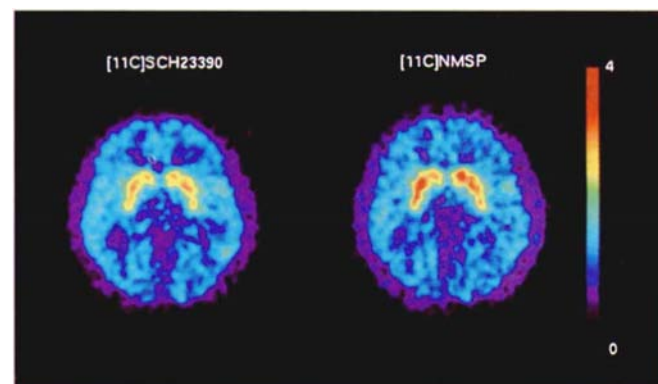


Figure 1 PET visualization of D1R (left image) and D2R (right image) in a control subject. The radioactivity distribution was obtained at the level of striatum, 14 to 40 min after injection of [¹¹C]SCH23390 and 34–60 min after [¹¹C]NMSP injection. Images were normalized with respect to cerebellar radioactivity. The *in vivo* labelling of the striatum and cortex by [¹¹C]SCH23390 (left) represents D1R, and [¹¹C]NMSP (right) binds predominantly to D2R in the striatum but to 5-HT₂ in the cortex.

Table 1 Mean values of D1R and D2R binding

	Controls (<i>n</i> = 18, 12*)	Schizophrenics		ANOVA <i>F</i> (<i>P</i>)	ANCOVA <i>F</i> (<i>P</i>)
		Drug-naïve (<i>n</i> = 10)	Drug-free (<i>n</i> = 7)		
D2 dopamine receptor (k_3); [^{11}C]NMSP					
Striatum	0.033 ± 0.007	0.036 ± 0.006	0.032 ± 0.004	1.02(0.38)	0.43(0.65)
D1 dopamine receptor (k_3/k_4); [^{11}C]SCH23390					
Striatum	1.29 ± 0.17	1.28 ± 0.15	1.20 ± 0.18	0.67(0.52)	0.47(0.63)
Prefrontal cortex	0.41 ± 0.06	0.35 ± 0.06†	0.33 ± 0.05†	6.15(0.005)	9.11(0.0008)
Anterior cingulate cortex	0.48 ± 0.08	0.42 ± 0.07	0.41 ± 0.07	3.40(0.046)	5.68(0.008)
Temporal cortex	0.46 ± 0.08	0.42 ± 0.06	0.38 ± 0.05	3.32(0.049)	4.35(0.022)
Occipital cortex	0.33 ± 0.05	0.33 ± 0.06	0.30 ± 0.05	1.04(0.37)	0.82(0.45)

k_3 , Association rate constant; k_4 , dissociation rate constant; k_3/k_4 , binding potential assumed to be equal to B_{max}/K_D . According to the one-way analysis of variance (ANOVA), a significant difference ($P < 0.0083 = 0.05/6$) was considered as significant to avoid type I errors in the multiplicity of statistical analyses among the groups was found only in the D1R of the PFC. Because we found an age effect on D1R and D2R binding, in agreement with earlier^{15,22}, an analysis of covariance (ANCOVA) with age as covariate was performed to test for group differences; significant differences in the PFC and anterior cingulate cortex were found.

* Twelve of the 18 control subjects underwent PET scans with [^{11}C]SCH23390 and [^{11}C]NMSP; 6 controls underwent PET scan with [^{11}C]SCH23390 only.

† Duncan's *post hoc* analysis after ANOVA showed that the mean values of D1R binding in the PFC for both patient groups were significantly smaller than for the control group ($P < 0.05$).

spiperone([^{11}C]NMSP), respectively, as the PET tracers (Fig. 1 and Methods). *In vivo* labelling of the striatum and cortex by [^{11}C]SCH23390 is thought to be due to binding to dopamine D1 receptors^{11,12}; [^{11}C]NMSP binds predominantly to dopamine D2 receptors in the striatum¹³.

We found no difference in striatal D2R among the three groups (Table 1 and Fig. 2a). This observation is in agreement with PET findings using different radioligands²⁻⁵ and fails to support the hypothesis that schizophrenia is related to an increased density of D2 receptors.

As for D1R, a post-mortem study showed an increased ratio of D2R/D1R in the striatum of schizophrenics¹⁴, and a preliminary PET study has reported that D1R binding in the putamen is lower in schizophrenics, suggesting that a low D1R binding in the putamen is lower in schizophrenics, suggesting that a low D1R density may result in a reduced activity of D1R to D2R regulated feedback systems in schizophrenia¹⁵. However, we failed to detect this difference either in the striatal D1R (Table 1 and Fig. 2b) or in the ratio of striatal D1R to striatal D2R (data not shown).

In contrast to the lack of difference in the striatal D1R and D2R, the mean values of D1R binding potential in the prefrontal cortex (PFC) for both patient groups were significantly lower than for the control group. (Table 1 and Fig. 2c). Because the binding potential is not critically dependent on cerebral blood flow¹⁶, such difference cannot be accounted for by its reduction. To exclude the effect of difference in the size of the regions of interest (ROIs), the number of pixels for each ROI was calculated and no significant difference was

found in any region (data not shown). Our magnetic resonance imaging (MRI) measurements (see Methods) revealed no difference in the PFC volume of schizophrenics and controls (data not shown). Reduction in D1R binding is therefore not associated with gross alterations in brain anatomy. Furthermore, the decreased D1R binding in patients who have never been treated with neuroleptics indicates that the D1R system in PFC may be involved in the disease process of schizophrenia itself.

Reduction of prefrontal D1R was related to the clinical ratings for negative symptoms (Fig. 3a) and poor performance in the Wisconsin Card Sorting Test (WCST; Fig. 3b) but was not correlated with age at onset or duration of illness.

Several lines of evidence have suggested that the PFC of schizophrenic patients may be hypodopaminergic and that this decreased mesocortical dopaminergic activity may induce poor performance in frontally mediated cognitive tasks and negative symptoms¹⁷: (1) local depletion of dopamine in the PFC of monkeys produced an impairment in spatial delayed alternation performance which was almost as severe as that caused by surgical ablation of the same area, and this behavioural deficit could be reversed by a dopamine agonist¹⁸; (2) a study of human plasma using homovanillic acid indicated that the negative symptoms of schizophrenic patients are associated with decreased dopaminergic function¹⁹; (3) administration of dopamine stimulant ameliorates cognitive performance on frontally mediated tasks and task-dependent activation of regional blood flow in the PFC²⁰ in schizophrenics; and (4) a PET study has revealed that an impaired cognitive-task-induced activation of the

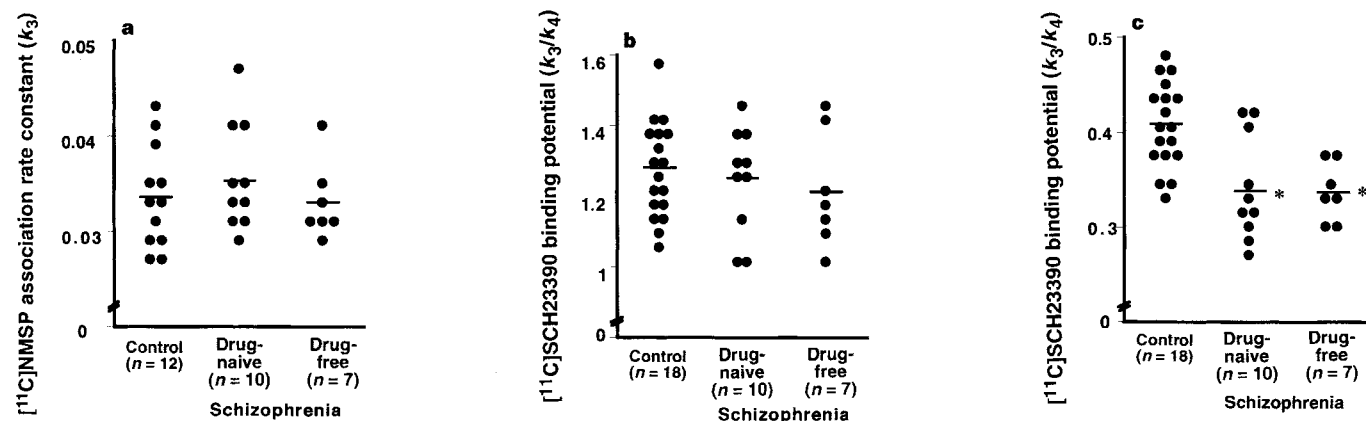


Figure 2 Distribution of D1R and D2R binding indices. Values were adjusted for age. **a**, Association rate constant (k_3) of striatal D2R: adjusted $k_3 = k_3 + 0.0006$ (age, 28.5 yr). **b**, Binding potential (k_3/k_4) of striatal D1R: adjusted $k_3/k_4 = k_3/k_4 + 0.16$ (age, 28.5 yr). **c**, Binding potential (k_3/k_4) of PFC D1R: adjusted

$k_3/k_4 = k_3/k_4 + 0.06$ (age, 28.5 yr). Asterisks, Duncan's *post hoc* analysis after ANOVA revealed that both patient groups had reduced D1R binding in the PFC compared with the control group ($P < 0.05$).

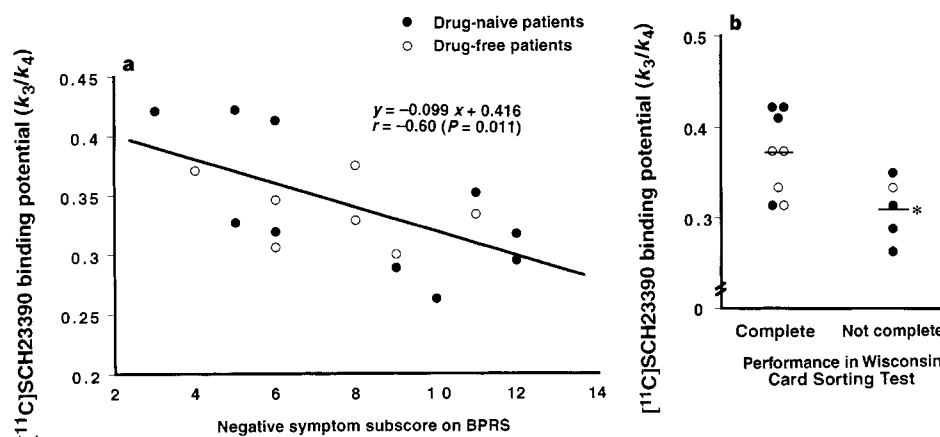


Figure 3 a, Prefrontal D1R binding potential adjusted for age negatively correlates with the negative symptom subscore (blunted affect, emotional withdrawal, motor retardation) on the Brief Psychiatric Rating Scale (BPRS). When restricted to drug-naïve patients only, the correlation was the same. Prefrontal D1R binding was not correlated with the BPRS total score or the positive-symptom subscore. **b**, Modified WCST^o was performed in 8 drug-naïve patients (filled circles) and 5 drug-treated patients (circles). Eight patients completed the required categories

within two steps ('complete' group) but the other 5 patients performed poorly and could not complete the categories ('not complete' group). Two-way ANCOVA (performance on WCST $\times$ history of medication) with age as covariate revealed that patients with poor performance showed decreased D1R binding in the PFC ($*F = 9.65, P = 0.014$). There was no significant interaction between performance and history of medication.

anterior cingulate cortex in schizophrenics can be significantly modulated by dopaminergic manipulation²¹.

Several findings indicate that the proposed hypodopaminergic activity in the PFC of schizophrenics may be due to hypofunction in the D1R system: (1) the human cortex contains 4 to 7 times more D1R than D2R (ref. 6); (2) administration of the D1 antagonist SCH23390 produced deficits in performance that were dose-dependent during a delayed-response task in monkeys⁷; (3) by combining iontophoretic analysis of dopamine receptors with single-cell recording delayed-response performance, the specific role of D1R in regulating neuronal activity associated with working memory in monkeys has been revealed⁸.

Here we have shown that D1R binding decreases in the PFC of schizophrenic patients and that this reduction is related to negative symptoms and poor performance in the WCST. Our findings indicate that this dysfunction in the D1R modulation mechanism in the PFC may contribute to negative symptoms and cognitive deficits in schizophrenia and suggest that selective D1 agonists may be helpful in treating the negative symptoms of schizophrenia. $\square$

Methods

Subjects. This study was approved by the Ethics and Radiation Safety Committee of the National Institute of Radiological Sciences and the Ethics Committee of Tokyo Medical and Dental university. We examined 17 patients aged 27.4 ± 5.9 yr (mean $\pm$ s.d.) and 18 controls aged 27.7 ± 5.6 yr. Ten of the patients (aged 26.1 ± 3.8 yr) were neuroleptic-naïve and 7 (aged 29.2 ± 8.1 yr) had been previously treated with neuroleptics but had been drug-free for at least two weeks (median 5 weeks, maximum 4 years) before the PET study. Written informed consent was obtained from all subjects. The patients all met the ICD-10 criteria for schizophrenia. The duration of illness ranged from 4 months to 18 years (average 5.5 yr).

PET experiments. Two PET experiments were performed on the same day with a four-ring, 7-slices PET system (PCT3600W40; Hitachi). Spatial resolution of the system is 8 mm on the plane and axial resolution is 12 mm. The subject were carefully positioned with the aid of a vertical laser line so that the lowest slice included the cerebellar hemispheres and was parallel to the orbitomeatal plane. The head was immobilized during scanning with a specially designed headholder and a plastic face mask. In the first PET run, a dose of 291.2–412.2 MBq of [¹¹C]SCH23390 was injected to study D1R in the striatum and cortex. The radioactivity was followed for 40 min immediately following injection. Two hours after the first examination, a dose of 226.4–563.1 MBq of [¹¹C]NMSP was injected to study D2R in the striatum and radioactivity monitored for 60 min.

Quantification of D1R and D2R binding. Regions of interest (ROIs) were defined blindly by Y.O. on the basis of a previously used²² semiautomatic method, which consists of two steps. First, areas in the brain were divided into several regions manually, with surrounding cerebrospinal fluid space and white matter on the PET image being referred to the brain atlas. Then the ROIs were defined by computer-controlled delineation of the percentage isocounter. The ROI for cerebellum was obtained in the lowest slice; ROIs for the striatum, PFC, anterior cingulate, temporal and occipital cortices in [¹¹C]SCH23390 images and the ROIs for the striatum in [¹¹C]NMSP images were obtained in the slice in which the striatum was most visible (Fig. 1). Based on a two-compartment model²³, the binding potentials of D1R, which are defined as the association rate constant (k_3) divided by the dissociation rate constant (k_4) and are assumed to be equal to $B_{\max}$ (maximum receptor density)/ K_D (dissociation constant) (ref. 16), were quantified in the striatum and PFC. Because the dissociation of [¹¹C]NMSP in the striatum is negligible, the association rate constant (k_3) was used for the D2R in the striatum²⁴. To compare the volume of the brain, the MRI scan was performed with 2-mm contiguous coronal sections. The measured values were summed and then divided by the cranium volume.

Received 15 August; accepted 4 December 1996.

- Seeman, P., Lee, T., Chau-Wong, M. & Wong, K. *Nature* **261**, 717–719 (1976).
- Hietala, J. *et al. Archs Gen. Psychiat.* **51**, 116–123 (1994).
- Martinet, J. L., Person-Magnan, P., Huret, J. D. & Mazoyer, B. *Am. J. Psychiat.* **147**, 44–50 (1990).
- Nordström, A. L., Farde, L., Eriksson, L. & Hallidin, C. *Psychiat. Res. Neuroimaging* **61**, 67–83 (1995).
- Farde, L. *et al. Archs Gen. Psychiat.* **47**, 213–219 (1990).
- Hall, H. *et al. Neuropsychopharmacology* **11**, 245–256 (1994).
- Sawaguchi, T. & Goldman-Rakic, P. S. *Science* **251**, 947–950 (1991).
- Williams, G. V. & Goldman-Rakic, P. S. *Nature* **376**, 572–575 (1995).
- Goldman-Rakic, P. S. *J. Neuropsychiat. Clin. Neurosci.* **6**, 348–357 (1994).
- Kashima, H. *Jpn J. Psychopharmacol.* **11**, 83–88 (1991).
- Farde, L., Hallidin, C., Stone, E. S. & Sedvall, G. *Psychopharmacology* **92**, 278–284 (1987).
- Anderson, P. & Gronvald, F. *Life Sci.* **38**, 1507–1514 (1986).
- Wong, D. F. *et al. Science* **226**, 1393–1396 (1984).
- Hess, E. J., Bracha, H. S., Kleinman, J. E. & Creese, I. *Life Sci.* **40**, 1487–1497 (1987).
- Sedvall, G. & Farde, L. *Lancet* **346**, 743–749 (1995).
- Mintun, M. A., Raichle, M. E., Kilbourn, M. R., Wooten, G. F. & Welch, M. J. *Ann. Neurol.* **15**, 217–227 (1984).
- Weinberger, D. R., Berman, K. F. & Zec, R. F. *Archs Gen. Psychiat.* **43**, 114–124 (1986).
- Brozoski, T., Brown, R., Rosvold, H. & Goldman-Rakic, P. S. *Science* **205**, 929–931 (1979).
- Davidson, M. & Davis, K. *Archs Gen. Psychiat.* **45**, 561–563 (1988).
- Daniel, D., Weinberger, D. & Jones, D. J. *Neurosci.* **11**, 1907–1917 (1991).
- Dolan, R. J. *et al. Nature* **378**, 180–182 (1995).
- Asahina, M. *et al. Acta Neurol. Scand.* **91**, 437–443 (1995).
- Suhara, T. *et al. Psychopharmacology* **103**, 41–45 (1991).
- Patlak, C. & Blasberg, R. J. *Cereb. Blood Flow Metab.* **5**, 584–590 (1985).

Acknowledgements. We thank F. Shishido, H. Mori, S. Tsung and J. Ji for their help. Supported by the PET project of the National Institute of Radiological Sciences of Japan and by a grant-in-aid for Scientific Research from the Ministry of Education of Japan.

Correspondence should be addressed to Y.O. (e-mail: okubo.psyc@med.tmd.ac.jp).

Bax suppresses tumorigenesis and stimulates apoptosis *in vivo*

Chaoying Yin*, C. Michael Knudson†, Stanley J. Korsmeyer† & Terry Van Dyke*

* Department of Biochemistry and Biophysics, Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine at Chapel Hill, Chapel Hill, North Carolina 27599, USA

† Howard Hughes Medical Institute and Division of Molecular Oncology, Washington University School of Medicine, St Louis, Missouri 63110, USA

The protein p53 is a key tumour-suppressor, as evidenced by its frequent inactivation in human cancers. Animal models have indicated that attenuation of p53-dependent cell death (apoptosis) can contribute to both the initiation and progression of cancer, but the molecular mechanisms are unknown. Although p53-mediated transcriptional activation is one possible explanation, none of the known p53-responsive genes has been shown to function in p53-dependent apoptosis. Here we test the role of the death-promoting gene *bax* in a transgenic mouse brain tumour, a model in which p53-mediated apoptosis attenuates tumour growth. Inactivation of p53 causes a dramatic acceleration of tumour growth owing to a reduction in apoptosis of over ninety

per cent¹. We show that p53-dependent expression of *bax* is induced in slow-growing apoptotic tumours. Moreover, tumour growth is accelerated and apoptosis drops by fifty per cent in Bax-deficient mice, indicating that it is required for a full p53-mediated response. To our knowledge this is the first demonstration that Bax acts as a tumour suppressor, and our findings indicate that Bax could be a component of the p53-mediated apoptotic response in this system.

Bax is a Bcl-2 related protein that promotes apoptosis in cultured cells². Mice deficient in Bax have selective hyperplasias and show resistance to certain apoptotic stimuli³. As p53 activity correlates with increased Bax expression^{4,5} and activates expression through p53-responsive elements in the *bax* promoter⁶, it has been proposed that Bax might act downstream of p53 in the induction of apoptosis. However, Bax-deficient thymocytes are not defective in p53-dependent DNA-damage-induced apoptosis³, as are p53-deficient thymocytes^{7,8}. Loss of this pathway in p53-null mice predisposes them to T-cell tumorigenesis⁹; a role for Bax in tumour suppression has not been established.

Because tissue-specific pathways are probably involved in tumorigenesis of different cell types, we examined the role of Bax in the transgenic(Tg)T₁₂₁ model of p53 tumour suppression. The cells targeted in these transgenic mice, choroid plexus (CP) epithelium, normally withdraw from the cell cycle within 2 weeks after birth and then remain in a resting state¹⁰. CP cells are induced to proliferate aberrantly by tissue-specific expression of T₁₂₁, a truncated SV40 T antigen that inactivates the retinoblastoma(Rb) proteins but not p53^{11,12}. Tumour growth is suppressed in these mice by induction of

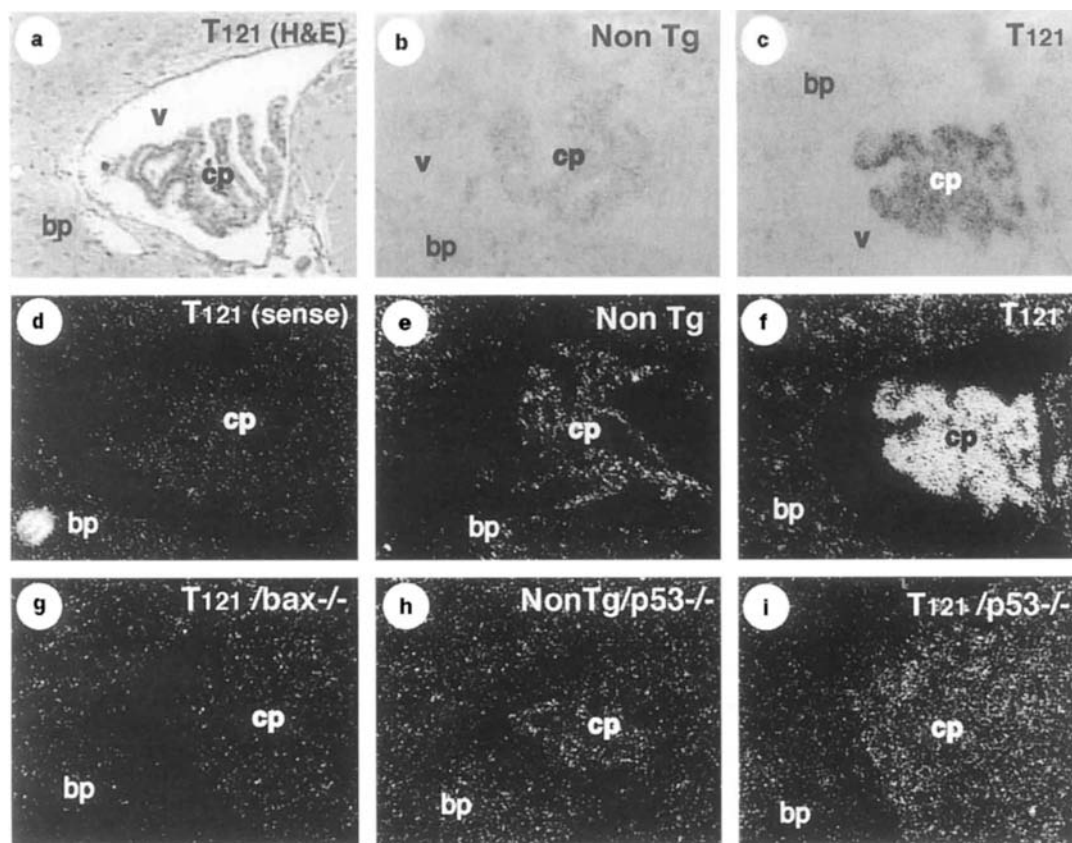


Figure 1 p53-dependent *bax* expression in apoptotic TgT₁₂₁ CP. A haematoxylin-and-eosin(H&E)-stained section (**a**) shows the brain region present in all panels. CP tissue (cp) lies within the ventricular space (v) which is surrounded by brain parenchyma (bp). Representative signals from *in situ* RNA hybridization with a *bax* antisense RNA probe are shown for non-Tg (**b** and **e**), TgT₁₂₁ (**c** and **f**), TgT₁₂₁*bax*^{-/-} (**g**), non-Tg *p53*^{-/-} (**h**) and TgT₁₂₁*p53*^{-/-} (**i**) mice. Hybridization of

TgT₁₂₁ sections with a control *bax* sense probe is shown in **d**. Bright-field (**a-c**) and dark-field (**d-i**) views are shown. Expression of *bax* is low in non-Tg CP but highly induced in TgT₁₂₁ CP (compare **b** with **c**, **e** with **f**). Most of this expression is p53-dependent as the level is greatly reduced in the CP of TgT₁₂₁*p53*^{-/-} mice (compare **f** and **i**). Quantification of signal density is shown in Table 1. Magnification, 60 ×.

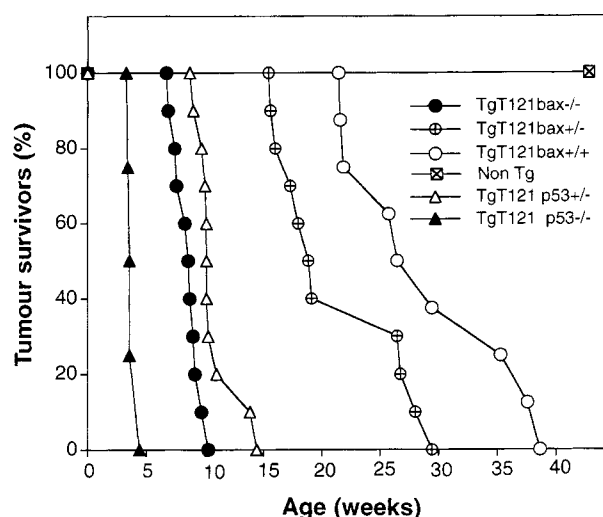


Figure 2 Bax is a tumour suppressor. Bax deficiency accelerates tumour growth, as indicated by a decrease in survival time, t_{50} . Survival of TgT₁₂₁ mice is shown relative to TgT₁₂₁ mice with a deficiency in either *p53* or *bax*, and to TgT₁₂₁ mice carrying a single inactive *p53* or *bax* allele. Survival times reflect the time of killing dictated by the presence of a severely bulged cranium. Mice generally survive only 1–2 days past this point. Histological examination confirmed the presence of tumour masses in all mice. For each curve, a single point represents a single mouse, except that at least 20 mice are represented by the non-Tg point (square with cross). TgT₁₂₁*bax*^{+/+} mice in this study were derived directly from the crosses used to generate mice with altered *bax* genotypes and showed no change in average survival time relative to other TgT₁₂₁ mice previously studied. The survival of mice with altered *p53* genotypes were previously reported¹.

p53-dependent apoptosis¹. To determine whether *bax* expression is induced in T₁₂₁-expressing cells, we compared levels of *bax* transcripts in the CP of normal, TgT₁₂₁ and TgT₁₂₁*p53*^{-/-} mice. Young mice (4 weeks old) were used to avoid any secondary effects of long-term tumour growth. Representative CP morphology at this stage is shown in Fig. 1a. *In situ* RNA hybridization showed a marked 8-fold induction of *bax* transcripts in TgT₁₂₁ CP compared with normal (Fig. 1: compare b with c (bright-field), and e with f (dark-field); Table 1). No hybridization above background was detected in the CP of TgT₁₂₁*bax*^{-/-} mice (Fig. 1g) or with a sense probe control (Fig. 1d). Induced *bax* expression was limited specifically to the CP; in brain regions outside CP that did not express T₁₂₁, *bax* expression was comparable in non-transgenic and in TgT₁₂₁ mice (Fig. 1, and data not shown). Induction of *bax* in the CP was *p53*-dependent, because *bax* transcript levels were much lower in the CP of TgT₁₂₁*p53*^{-/-} mice (Fig. 1: compare f and i; Table 1). A comparably small amount of *bax* RNA was expressed in the absence of *p53* in both non-transgenic (Fig. 1h) and TgT₁₂₁ CP (Fig. 1i and Table 1). Hence, induction of *p53*-dependent Bax expression accompanies induction of *p53*-dependent apoptosis in this system.

To determine whether Bax is involved in tumour suppression, we examined the development of T₁₂₁-induced tumours in a Bax-deficient background³. The rate of tumour development was assessed in TgT₁₂₁*bax*^{-/-} mice and TgT₁₂₁ littermates as a function of survival (Fig. 2). As previously shown¹¹, TgT₁₂₁ mice developed slow-growing tumours, with a 50% survival (t_{50}) of 26 weeks (Fig. 2). When both *bax* alleles were inactivated, tumour growth rate increased significantly. Figure 2 shows that although both *p53* alleles were intact, t_{50} for TgT₁₂₁*bax*^{-/-} mice was only 8.5 weeks, and was 3.5 weeks for TgT₁₂₁ mice deficient in *p53* (ref. 1). Histological analysis of tumour tissue from TgT₁₂₁, TgT₁₂₁*bax*^{-/-} and TgT₁₂₁*p53*^{-/-} mice revealed uniformly neoplastic growth of similar morphology in all animals (Fig. 3a, c, d), although much more time was required in TgT₁₂₁ mice (Fig. 2, and compare relative mass sizes in Fig. 3a with c, d). These results show that Bax inactivation leads to rapid tumour growth and indicate that Bax is a tumour suppressor.

We have previously shown that attenuated tumour growth in the presence of wild-type *p53* correlates with extensive apoptosis. The apoptotic index (AI) of TgT₁₂₁ CP averaged 7.3%, whereas that of TgT₁₂₁*p53*^{-/-} mice was <1%, constituting >90% reduction¹. If Bax is involved in *p53*-mediated tumour suppression, the CP of TgT₁₂₁*bax*^{-/-} mice should also show a marked reduction in the CP AI. Furthermore, if AI is inversely correlated with tumour growth rate, an AI intermediate between that of TgT₁₂₁ and TgT₁₂₁*p53*^{-/-} mice would be expected as tumour growth rate in TgT₁₂₁*bax*^{-/-} mice was intermediate. We therefore assayed brain CP from several TgT₁₂₁ and TgT₁₂₁*bax*^{-/-} mice *in situ* for apoptosis using the TUNEL assay¹³ (see Methods; Fig. 3e, g, respectively). As predicted, analysis of three independent mouse sets showed AI (Fig. 4a) to be significantly reduced by an average of 49% in TgT₁₂₁*bax*^{-/-} mice compared with their TgT₁₂₁ littermates ($4.5 \pm 0.9\%$ compared to $8.8 \pm 2.3\%$, respectively; Fig. 4b, left). AI was similar in mice of different ages (Fig. 4a), indicating that the effect was due to Bax deficiency and not to other parameters of tumour growth. No apoptotic cells were detected in normal CP from non-transgenic or from *bax*^{-/-} mice (data not shown). As AI is reduced by >90% in *p53* deficiency¹, these data indicate that about half of T₁₂₁-induced *p53*-dependent apoptosis is mediated through Bax. This is consistent with the observation that T₁₂₁ induces substantial *p53*-dependent *bax* expression in apoptotic CP (Fig. 1; Table 1). *p53*-independent apoptosis (only ~10% of total) might also require Bax as a small amount of *bax* expression persists in the absence of *p53* (Fig. 1i; Table 1). Because an intermediate reduction in AI leads to an intermediate acceleration of tumour growth, there is a direct relation between the extent of apoptosis and tumour suppression (Fig. 4b).

Bax, like other Bcl-2 family proteins, can form both homo- and heterodimers, suggesting that one possible mechanism for its action is to titrate one or more of the apoptosis-inhibiting members². In cultured cells, the amount of apoptosis can vary depending on the relative amounts of Bax and Bcl-2 present². As inactivation of one *bax* allele could effectively reduce the amount of Bax in cells, TgT₁₂₁*bax*^{+/-} mice might undergo some acceleration of CP tumour growth as a result of reduced apoptosis. We therefore analysed AI and tumour growth in TgT₁₂₁*bax*^{+/-} mice. Three young mice (Fig. 4a) showed an average reduction in AI of 23%, which compares to a 49% reduction in TgT₁₂₁*bax*^{-/-} mice (Fig. 4b, left). Tumour growth rate was accelerated, as indicated by a reduction in t_{50} from 26 to 18.5 weeks. Such a decrease in survival time could also result from rapid focal outgrowth of tumours with secondary genetic lesions, as occurs in all TgT₁₂₁*p53*^{+/-} mice¹. To test this in TgT₁₂₁*bax*^{+/-} mice, we examined the CP at 3-week

Table 1 Relative *bax* RNA levels in choroid plexus

P53 status	Probe	Non-Tg	T ₁₂₁	T ₁₂₁ <i>bax</i> ^{-/-}
<i>p53</i> ^{+/+}	Sense	ND	1.0	ND
<i>p53</i> ^{-/-}	Antisense	45.5	346.0	2.0
<i>p53</i> ^{-/-}	Antisense	14.2	16.5	ND

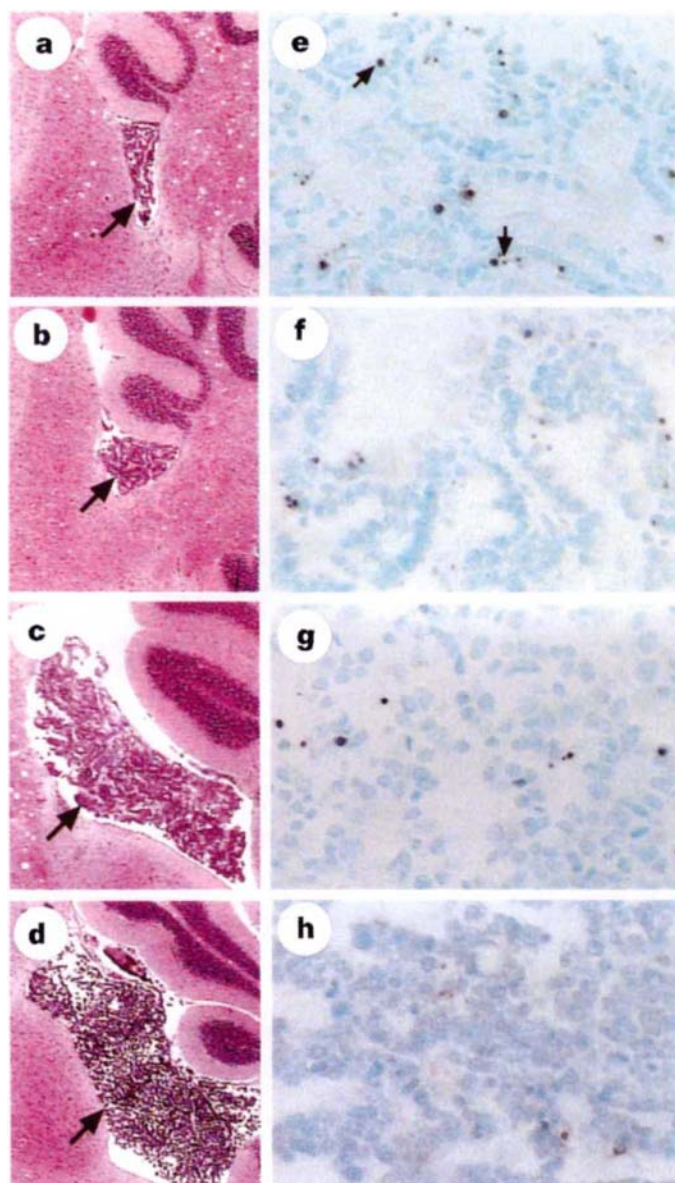


Figure 3 Tumour growth is accelerated and apoptosis is diminished in $TgT_{121}bax^{-/-}$ mice. Morphology of CP masses (arrows) is shown for $TgT_{121}bax^{-/-}$ (b) and $TgT_{121}p53^{-/-}$ (c) mice relative to TgT_{121} (a) CP. All mice were 6 weeks old, a time that is almost terminal for $TgT_{121}bax^{-/-}$ mice (Fig. 2). A terminal $TgT_{121}p53^{-/-}$ (3.5 weeks) brain is shown in d. Sections (6 microns) were stained with haematoxylin and eosin. Representative fields showing apoptosis as detected by TUNEL is shown for the same samples: TgT_{121} (e), $TgT_{121}bax^{-/-}$ (f), $TgT_{121}bax^{-/-}$ (g) and $TgT_{121}p53^{-/-}$ (h) CP. Arrows in e point to a few representative apoptotic cells. Note the inverse correlation between apoptosis levels and tumour growth by comparing apoptosis and CP size relative to age. Average AIs for all samples were quantified and are shown in Fig. 4.

intervals from 3 to 18 weeks for morphological signs of focally aggressive growth. The frequency of focal tumour progression was similar to that in $TgT_{121}bax^{+/+}$ mice (data not shown), consistent with a reduced survival time resulting from the overall reduction in apoptosis (Fig. 4b). This indicates a distinction between Bax and p53 and could reflect an increased contribution to tumour progression from loss of the p53 allele as a result of genetic instability¹⁴⁻¹⁶.

Our results show that Bax can function as a tumour suppressor and are the first indication *in vivo* that a p53-regulated gene is involved in p53-dependent apoptosis. We have not been able to

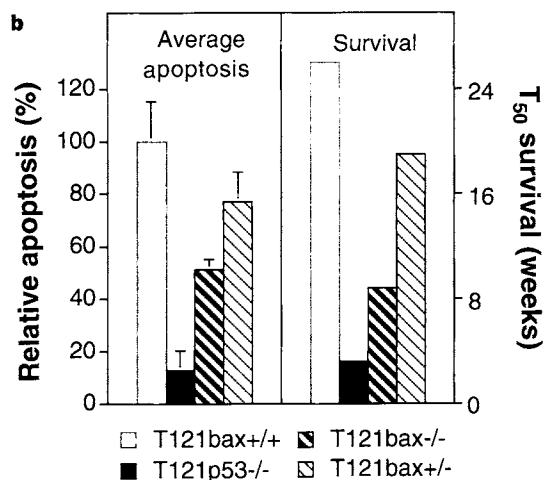
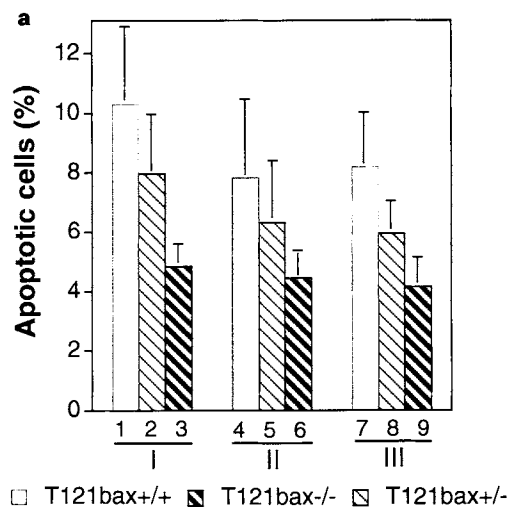


Figure 4 Apoptosis correlates inversely with tumour growth rate. In a, AIs were quantified for three independent experimental sets (I-III) of mice and absolute values are given. In each set TgT_{121} (1,4,7), $TgT_{121}bax^{-/-}$ (2,5,8), and $TgT_{121}bax^{+/+}$ (3,6,9) littermates were examined. Mice within each set were 3.5 weeks (II), 4 weeks (III) and 6 weeks (I) of age. Standard errors represent the variation in quantifying 10 independent fields of view for each mouse. In b, left, average AI is presented relative to the TgT_{121} level for $TgT_{121}p53^{-/-}$, $TgT_{121}bax^{-/-}$ and $TgT_{121}bax^{+/+}$ mice. Standard errors reflect variations in average values between individual mice. Mean survival time indicative of tumour growth rate is shown in the right panel and was derived from the survival in Fig. 2.

determine whether Bax is immediately downstream of p53, although p53-dependent induction of *bax* in T_{121} -expressing cells is consistent with Bax as a downstream effector. Although our *in vivo* system cannot demonstrate direct induction by p53, it shows how important Bax is for complete tumour suppression by p53. Our results indicate that Bax is required for half of the p53-dependent apoptosis. McCurrach *et al.* also recently showed that half of p53-dependent apoptosis induced by adenovirus E1A in cultured fibroblasts requires Bax¹⁷. In these systems, the residual p53-dependent apoptosis not requiring Bax (~50%) could result from

compensation by other factor(s), regulation of other effectors by p53, or p53 induction of apoptosis by a transactivation-independent mechanism. In this regard, induction of full apoptosis in cultured cells requires synergism between transcription-dependent and -independent p53 functions¹⁸. Our experiments have uncovered an interesting specificity in p53 apoptotic mechanisms. In thymocytes, p53 is clearly required for DNA-damage-induced apoptosis^{7,8}, yet Bax is dispensable for this process³ and cannot rescue the response in p53-deficient thymocytes¹⁹. As Bax is required in CP for optimal p53-dependent apoptosis, there is a distinction between the tumour-suppression mechanisms employed: such differences could be due to cell type or the induction stimulus. It will be important to establish whether the progression of any human cancers also correlate with Bax inactivation. Given the role for Bax in p53-mediated tumour suppression shown here, those tumours that retain wild-type p53 alleles will be of particular interest. □

Methods

Generation of Bax-deficient TgT₁₂₁ mice. TgT₁₂₁ mice have been described¹. T₁₂₁ encodes the first 121 amino acids of SV40 T antigen followed by 11 missense residues¹². *bax*^{+/-} mice were generated by backcrossing *bax*^{-/-} mice³ for 4 generations with C₃H mice. Male *bax*^{+/-} mice were crossed for 2 generations with TgT₁₂₁ mice to generate TgT₁₂₁*bax*^{-/-} mice. Mice were identified by PCR analysis of tail DNA^{3,20}.

Histology and apoptosis assays. Brains were fixed in 10% formalin and embedded in paraffin, and sections (6 microns) were stained with a haematoxylin and eosin as described¹. To examine tumours, brains were sectioned from ten successive layers at 100-micron intervals. Apoptotic cells were detected in sections using the TUNEL (TdT-mediated dUTP nick end-labelling) assay¹.

In situ RNA hybridization. Sections (6 microns) were deparaffinized and hybridized as described²¹ with the following changes: proteinase K (20 µg ml⁻¹) incubation in 50 mM Tris (pH 7.6), 5 mM EDTA was for 7.5 min at room temperature, followed by fixation in 4% paraformaldehyde for 5 min. For probe templates, an *Apal*/*Bam*HI fragment of the mouse *bax* cDNA (exons 2–5) was inserted into *Apal*/*Bam*HI-digested pBS-SK(+/-). The antisense probe was generated by T3 transcription of a *Bsp*120I-linearized template and the sense probe by T7 transcription of a *Bam*HI-linearized template. Probes (5 × 10⁴ c.p.m. µl⁻¹) were radiolabelled using [α -³⁵S]UTP. Autoradiography was at 4 °C for 3 weeks. Sections were counterstained with 0.2% toluidine blue. Signal densities were quantified using NIH Image 1.58. The threshold was set so that sense probe signals were just detectable. Means were determined from 10 random areas of CP for each sample.

Received 13 November; accepted 30 December 1996.

1. Symonds, H. *et al.* *Cell* **78**, 703–711 (1994).
2. Oltvai, Z., Millman, C. & Korsmeyer, S. *Cell* **74**, 609–619 (1993).
3. Knudson, C. M., Tung, K. S. K., Tourtellotte, W. G., Brown, G. A. J. & Korsmeyer, S. J. *Science* **270**, 96–99 (1995).
4. Miyashita, T. *et al.* *Oncogene* **9**, 1799–1805 (1994).
5. Han, J. H. *et al.* *Genes Dev.* **10**, 461–477 (1996).
6. Miyashita, T. & Reed, J. C. *Cell* **80**, 293–299 (1995).
7. Clarke, A. R. *et al.* *Nature* **362**, 849–852 (1993).
8. Lowe, S. W., Schmitt, E. M., Smith, S. W., Osborne, B. A. & Jacks, T. *Nature* **362**, 847–849 (1993).
9. Donehower, L. A. *et al.* *Nature* **356**, 215–221 (1992).
10. Van Dyke, T. A. in *Molecular Genetics of Nervous System Tumors* (eds Levine, A. J. & Schmelk, H. H.) 287–301 (Wiley-Liss, New York, 1993).
11. Saenz-Robles, M. T., Symonds, H., Chen, J. & Van Dyke, T. *Mol. Cell. Biol.* **14**, 2686–2698 (1994).
12. Pipas, J. M., Peden, K. W. & Nathans, D. *Mol. Cell. Biol.* **3**, 203–213 (1983).
13. Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. *J. Cell Biol.* **119**, 493–501 (1992).
14. Lane, D. P. *Nature* **358**, 15–16 (1992).
15. Yin, Y., Tainsky, M. A., Bischoff, F. Z., Strong, L. C. & Wahl, G. M. *Cell* **70**, 937–948 (1992).
16. Livingstone, L. R. *et al.* *Cell* **70**, 923–935 (1992).
17. McCurrach, M. E., Connor, T. M. F., Knudson, C. M., Korsmeyer, S. E. & Lowe, S. W. *Proc. Natl Acad. Sci. USA* (submitted).
18. Chen, X., Ko, L. J., Jayaraman, L. & Prives, C. *Gene Dev.* **10**, 2438–2451 (1996).
19. Brady, H. J. M., Salomons, G. S., Bobeldijk, R. C. & Berns, A. J. M. *EMBO J.* **15**, 1221–1230 (1996).
20. Symonds, H. S., McCarthy, S. A., Chen, J., Pipas, J. M. & Van Dyke, T. *Mol. Cell. Biol.* **13**, 3255–3265 (1993).
21. Yan, C., Costa, R. H., Darnell, J. E., Chen, J. D. & Van Dyke, T. A. *EMBO J.* **9**, 869–878 (1990).

Acknowledgements. We thank H. Pan for advice on *in situ* hybridization and the other members of T.V.D.'s laboratory for help, and K. Berns and T. Bartolotta for histology sample preparation.

Correspondence and requests for materials should be addressed to T.V.D. (e-mail: tavandy@email.unc.edu).

A new class of membrane-bound chemokine with a CX₃C motif

J. Fernando Bazan*, Kevin B. Bacon†, Gary Hardiman*, Wei Wang†, Ken Soo†, Devora Rossi†, David R. Greaves‡, Albert Zlotnik† & Thomas J. Schall†

Departments of * Molecular Biology and † Immunology, DNAX Research Institute, 901 California Avenue, Palo Alto, California 94304, USA

‡ Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Chemokines direct the trafficking of white blood cells in immune surveillance, playing a key role in inflammatory and infectious diseases such as AIDS^{1–5}. All chemokines studied so far are secreted proteins of relative molecular mass ~7K–15K and fall into three families that are defined by a cysteine signature motif: CXC, CC and C (refs 3, 6, 7), where C is a cysteine and X any amino-acid residue. We report here the identification and characterization of a fourth human chemokine type, derived from non-haemopoietic cells and bearing a new CX₃C fingerprint. Unlike other chemokine types, the polypeptide chain of the human CX₃C chemokine is predicted to be part of a 373-amino-acid protein that carries the chemokine domain on top of an extended mucin-like stalk. This molecule can exist in two forms: either membrane-anchored or as a shed 95K glycoprotein. The soluble CX₃C chemokine has potent chemoattractant activity for T cells and monocytes, and the cell-surface-bound protein, which is induced on activated primary endothelial cells, promotes strong adhesion of those leukocytes. The structure, biochemical features, tissue distribution and chromosomal localization of CX₃C chemokine all indicate that it represents a unique class of chemokine that may constitute part of the molecular control of leukocyte traffic at the endothelium.

We searched for new chemokine and chemokine-like sequences from several non-haemopoietic cells and tissues using selected random sequencing, hybridization with chemokine probes, and by sequence and pattern searches⁸ of the expressed-sequence tag (EST) database at the National Center for Biotechnology Information (NCBI). A screen with the lone C-type chemokine lymphotactin revealed a new chemokine-like sequence with a CX₃C spacing of the hallmark cysteine motif. One of these ESTs (Genbank accession number H06104; corresponding to clone 44145 from the IMAGE consortium¹⁰) was sequenced. The resultant 1,635-base-pair (bp) complementary DNA clone, ending in a typical poly(A) stretch, as well as a splice variant (IMAGE clone 159419; Genbank accession number H14940) contained a common open reading frame of 1,194 bp that encoded a 397-amino-acid protein chain (Fig. 1). A predicted signal peptide accounts for the N-terminal 24 amino acids; directed sequence comparison with C, CC and CXC chemokine sequences^{6,7} showed that a chemokine module comprises the first 76 amino acids of the presumptive mature protein chain. But, unlike all other chemokine sequences that terminate at the end of this domain, the CX₃C chemokine protein continues for a further 241 residues which are rich in Gly, Pro, Ser and Thr. An uninterrupted stretch of 18 hydrophobic residues (amino acids 318–336) are likely to form a transmembrane link to a 37-amino-acid intracellular tail domain (Fig. 1). The predicted architecture of this molecule is therefore of a novel CX₃C chemokine fold resting on top of a long stalk, tethered to the surface of the cell membrane.

Alignment of the chemokine domain showed a closer relation with CC and C chemokines than with CXC (not shown). The stalk segment gave consistent database matches⁸ to regions of mucin

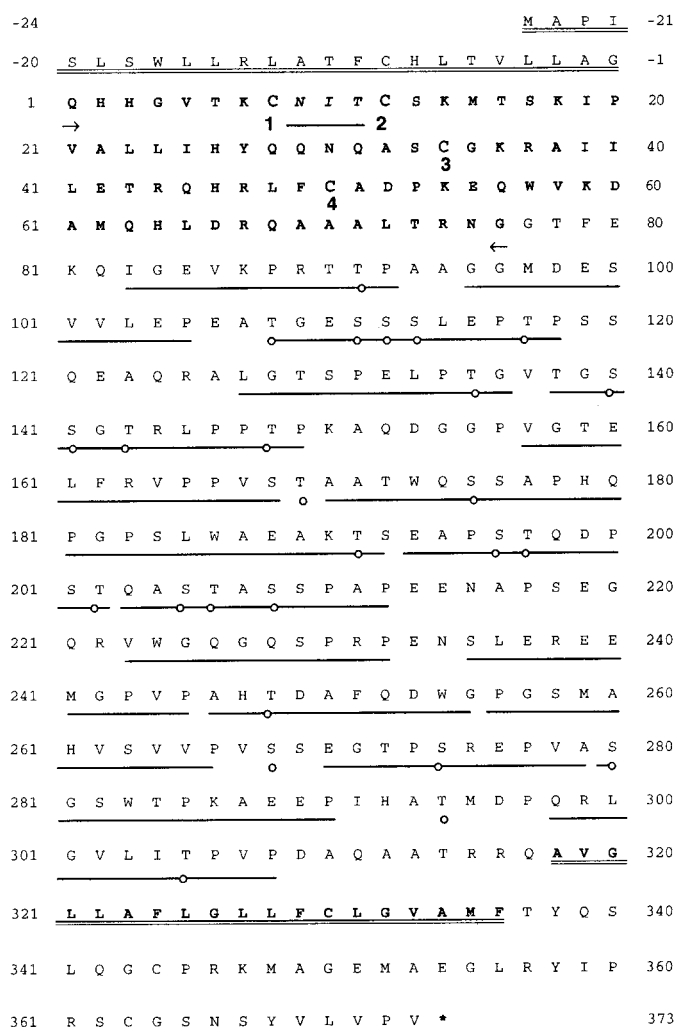


Figure 1 Amino-acid sequence and predicted structure of human CX₃C (fractalkine). The cDNA sequences (Genbank accession U84487) contain an open reading frame of 1,191 bp predicting a 397-amino-acid protein chain. The predicted signal peptide is indicated by a double underline and numbered -24 to -1; the N-terminal chemokine module (residues 1-76 of the mature sequence, in bold; boundaries indicated by arrows) features the signature Cys residues numbered 1-4, with the unique three-residue insertion in italics (and underlined to denote the N-X-T-N-glycosylation motif). The stalk region (amino acids 77-317) contains 17 degenerate mucin-like repeats in a tandem array; these are singly underlined, with predicted O-glycosylated Ser and Thr residues^{11,12} indicated by circles. The transmembrane segment (residues 318-336) is in bold and doubly underlined; the intracellular domain extends from amino acids 337-373. Full nucleotide sequences can be found in Supplementary Information.

chains which were defined by their extensive O-glycosylation of Ser and Thr residues located in short tandem repeats of varying length and by their extended backbone structure¹¹. The CX₃C chemokine stalk sequence contains 17 imperfect repeats of 10-12 amino acids with a loose XGTSX₃-SPTP motif (Fig. 1). Many of the component Ser and Thr residues of this repeat are reliably predicted to be O-glycosylated by a neural network algorithm¹². The mucin-like framework for the CX₃C chemokine stalk suggests that it may act as a tall, negatively charged projection from the cell surface, akin to transmembrane mucins with globular domains fused to their glycosylated stalks, such as the immunoglobulin modules that crown the endothelial mucin MAdCAM-1 or the leukocyte mucin CD96 (refs 11, 13). In addition, the elongated CX₃C chemokine stalk is similar in its structure and composition to syndecan-type

transmembrane proteoglycans¹⁴, but lacks any heparan-attachment glycanation motifs. CX₃C chemokine does, however, possess a membrane-proximal dibasic motif (Thr-Arg-Arg-Gln at amino acids 314-317), positioned similarly to a dibasic cleavage site in syndecans¹⁴.

To examine the tissue distribution of CX₃C chemokine messenger mRNA, a full-length cDNA was used to probe northern blots containing poly(A)⁺ RNA from several organs. Most cells expressed one or two mRNA species: the one of 3.5 kilobases (kb) was most common, but another of ~2.5 kb was also seen in colon, prostate, heart and skeletal muscle (Fig. 2a). Heart and skeletal muscle also expressed smaller mRNA species, and skeletal muscle alone had a larger mRNA as well. Little or no CX₃C chemokine mRNA was seen in peripheral blood (Fig. 2a). This is unlike most chemokines, which are abundantly expressed in leukocytes, the major source used for cloning these molecules¹⁻³. Further analysis of CX₃C chemokine expression in the cDNA from sixteen separate libraries derived from resting and activated peripheral blood mononuclear cells (PBMC), T cell, B cell and natural killer (NK) cell clones proved that expression of CX₃C chemokine was only scant in these cell types, despite being strong for CC and CXC chemokines (results not shown).

We analysed a panel of human/rodent somatic-cell hybrids to localize the human CX₃C chemokine gene. Initial mapping using the polymerase chain reaction (PCR) and several sets of specific primers on genomic DNA from the somatic-cell hybrids consistently yielded a PCR product in only those hybrids containing human chromosome 16. We confirmed this by testing a panel of four independently derived human/mouse somatic-cell hybrids, each of which contained human chromosome 16. All of the independent hybrids consistently yielded a PCR product with a variety of different CX₃C chemokine primer pairs, an example of which is shown in Fig. 2b. Genomic DNA collected from the parental mouse cell was negative (Fig. 2b), whereas positive controls using primers against a known human chromosome-16-encoded gene, α -globulin-14 (ref. 15), confirmed specificity (Fig. 2b). These data indicate that the CX₃C chemokine gene is in a different location from the other human chemokine genes localized so far.

As the endothelium is a major site of chemokine action in leukocyte trafficking, we analysed primary human pulmonary aortic endothelial cells (HPAEC) and human umbilical vein endothelial cells (HUVEC) by northern blotting. We detected little or no CX₃C chemokine mRNA in the unactivated state but found that expression increased markedly after stimulation of the cells by tumour-necrosis factor (TNF) or interleukin (IL)-1 (Fig. 2c, left panel). The pattern of mRNA expression paralleled protein expression on the surface of the endothelial cells. A polyclonal antibody against the carboxyl-terminal portion of the CX₃C chemokine module stained consistently on the surface of activated but not unactivated HUVEC and HPAEC (Fig. 2c, right panel). These results confirm that the CX₃C chemokine can exist as a cell-surface-bound protein that is induced on endothelial cells after activation by inflammatory cytokines.

To investigate the properties of CX₃C chemokine, its cDNA was cloned into a mammalian expression vector and transfected into the human embryonic kidney 293 (HEK293) cell line using three plasmids. The first was a truncated CX₃C chemokine cDNA (nucleotides 1-700) encoding a predicted protein of about 180 amino acids, including the chemokine module and about half of the predicted mucin-like stalk, but devoid of the transmembrane-coding segment. This construct was used to obtain a soluble molecule secreted into the HEK293 culture supernatant. The second construct was a 'full-length' (1.7 kb) CX₃C chemokine cDNA encoding all of the predicted open reading frame, including the transmembrane- and intracellular-domain-coding regions. This construct was designed to gauge the properties of the 'native' form of the CX₃C chemokine. The third construct was a control expres-

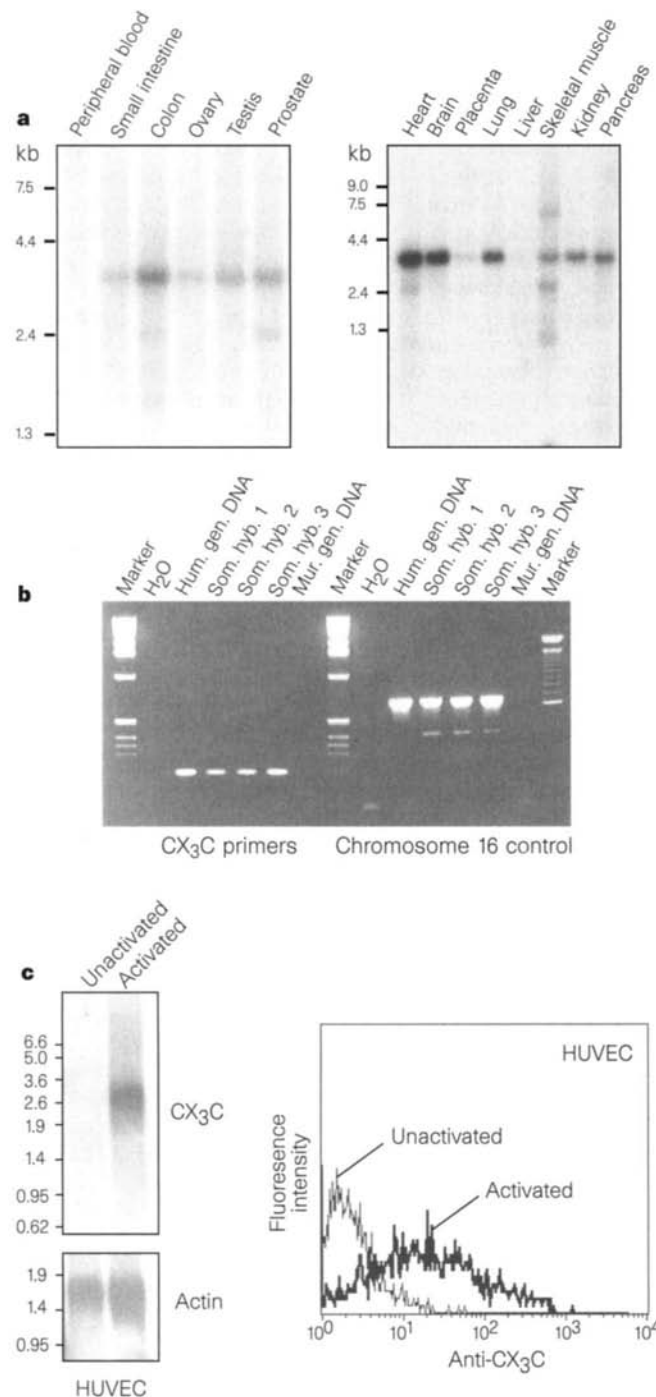


Figure 2 Expression patterns and chromosomal location of human fractalkine. **a**, Northern blots of poly(A)⁺ RNA showing CX₃C chemokine expression in mRNA from a variety of organs. **b**, Chromosomal localization of human CX₃C chemokine gene determined by PCR analysis of genomic DNA from a panel of rodent/human somatic-cell hybrid lines (som.hyb.1, som.hyb.3). Hum. gen. DNA, genomic DNA isolated from human fibroblasts; mur. gen. DNA, mouse genomic DNA isolated from MEL cells. Left half of the gel shows amplification in the presence of CX₃C chemokine-specific primers; amplification using primers specific for the human α -globin locus is shown on the right. **c**, CX₃C chemokine RNA and protein expression in human endothelial cells. Left, northern blot of CX₃C chemokine mRNA in HUVEC cells before and after treatment with either TNF or IL-1 ('unactivated' and 'activated'); actin probe is shown as a positive control. Right, protein expression by FACScan using an anti-CX₃C chemokine polyclonal antibody. Control purified IgG from the same animal did not stain (not shown). Almost identical results for both northern and FACScan analyses were obtained from several different HUVEC donors as well as from several HPAEC cell preparations.

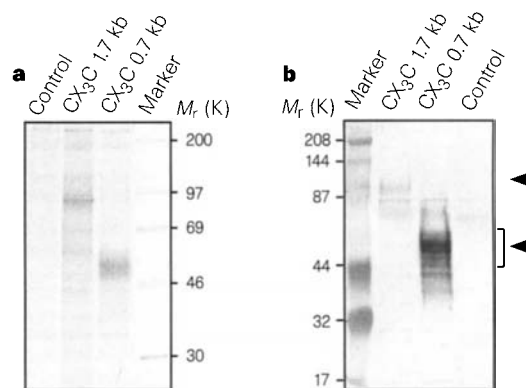


Figure 3 Secretion of recombinant fractalkine. **a**, Biosynthetically labelled supernatants of HEK293 cells transfected with expression plasmids encoding truncated CX₃C chemokine (CX₃C, 0.7 kb) or the full-length chemokine (CX₃C, 1.7 kb), or plasmid without insert ('Con') fractionated on 15% tricine-Tris SDS-PAGE. **b**, Western blot of CX₃C chemokine using a polyclonal antibody recognizing a peptide sequence in the chemokine module. Positions of full-length and truncated CX₃C chemokine proteins are indicated by arrows.

sion vector lacking a cDNA insert, designed to show the background pattern of protein expression from transfected HEK293 cells. Separate cultures of HEK293 cells were transfected with each of the three constructs, metabolically labelled with ³⁵S-methionine and cysteine, and the supernatants from the cell cultures were collected for analysis.

HEK293 cells transfected with the 0.7-kb construct gave a prominent and diffuse protein species of relative molecular mass (M_r) ~47K–55K in the cell culture supernatant; this protein was not present in the supernatant of control cells transfected with expression vector alone (Fig. 3a). We also detected a large protein species of M_r ~95K in the supernatants of the 1.7-kb full-length CX₃C chemokine cDNA transfectants that was not evident in the supernatants of the other cells (Fig. 3a). So, in addition to the truncated form of CX₃C chemokine (encoded by the 0.7-kb cDNA construct), at least some portion of the 'native' protein (encoded by the 1.7-kb cDNA) was being shed into the medium of transfected cells. Western blotting was used to confirm that the bands represented authentic CX₃C chemokine. The anti-CX₃C chemokine antibody stained diffuse bands centred at M_r 95K (top arrow in Fig. 3b) and 50K (bracketed region and lower arrow in Fig. 3b) in the supernatants of cells transfected with the 1.7- and 0.7-kb cDNAs, respectively, but not in control cells. These bands therefore represent recombinant secreted CX₃C chemokine protein. Glycosidase treatment reduced the apparent M_r of both the 47K–55K and 95K forms of the shed CX₃C chemokine (data not shown). These data indicate that CX₃C chemokine is encoded as a transmembrane molecule, but can be released into the supernatants of transfected cells as a 95K glycoprotein, possibly by proteolysis at the syndecan-like dibasic cleavage site proximal to the membrane.

We assessed the ability of the soluble 95K form of the CX₃C chemokine protein to induce the directed migration of leukocytes *in vitro* in a series of microchemotaxis assays. In each test, recombinant CX₃C chemokine collected in the medium of transfected cells was considerably more potent than the positive control, the CC chemokine RANTES, in the attraction of monocytes and T cells (Fig. 4a); supernatants from control transfectants had no attractant activity on their own. CX₃C chemokine was inactive on neutrophils (Fig. 4a), although those cells responded well to the CX₃C chemokine IL-8. Thus, like CC chemokines, CX₃C appears to be a potent attractant for monocytes and lymphocytes but not neutrophils. The

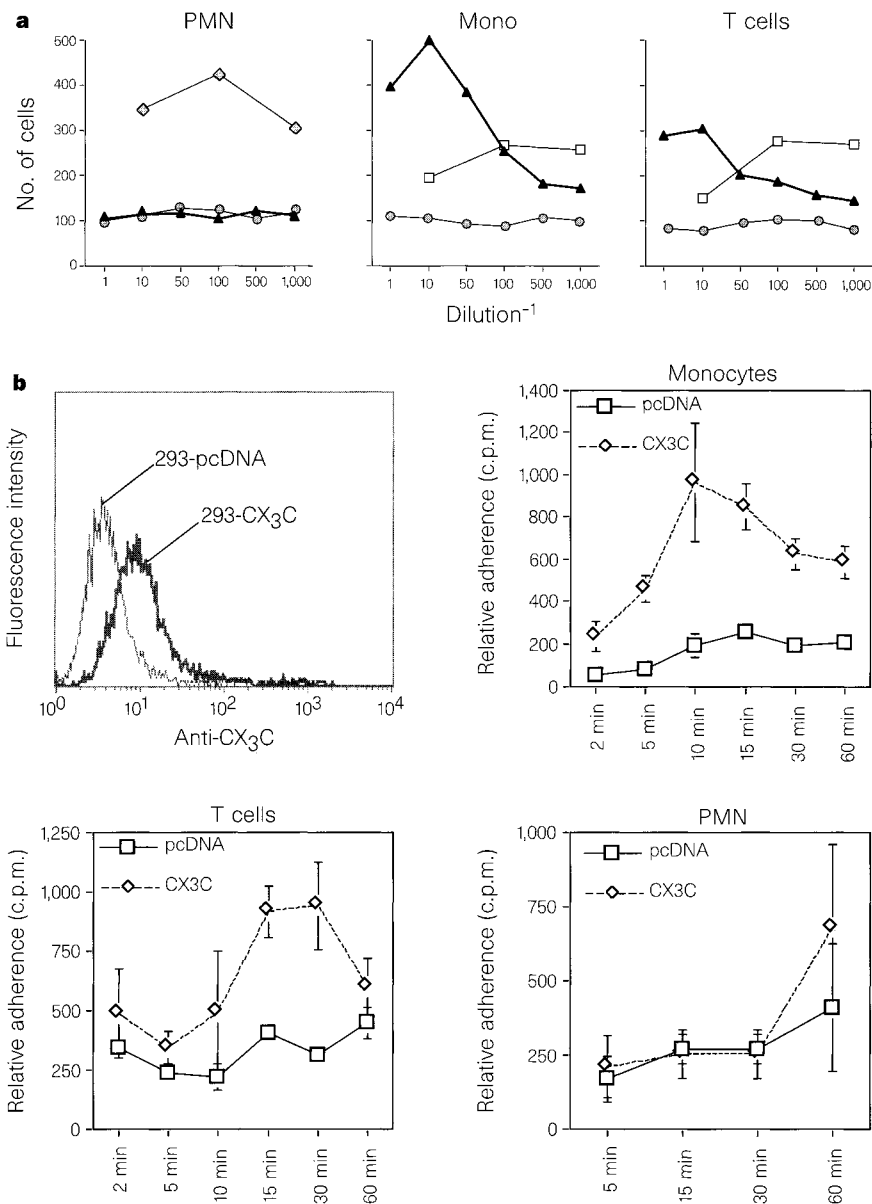


Figure 4 Leukocyte chemotactic and proadhesive activities of soluble and membrane-bound fractalkine. **a**, Results of microchemotaxis assays on peripheral blood neutrophils (PMN), monocytes (Mono) and T cells. Points represent serial dilutions of 'full length' (~95K) CX₃C chemokine glycoprotein in the supernatants of transfected cells (triangles); supernatants from control transfectants (circles) show background migration. Positive controls were purified recombinant human IL-8 used at 10⁻¹⁰ M, 10⁻⁹ M or 10⁻⁸ M (diamonds) on neutrophils, and purified recombinant human RANTES at the same concentration for monocytes and at 10⁻¹¹ M, 10⁻¹⁰ M or 10⁻⁹ M for lymphocytes (filled squares). Migration is

expressed as cell number per five high-power (× 400) fields, with duplicate wells being counted for each of four experiments. **b**, Adhesion of purified peripheral blood leukocyte populations to HEK293 cells stably expressing cell-surface CX₃C chemokine. Top left, levels of CX₃C chemokine on transfectants (as determined by FACSscan using the anti-CX₃C chemokine antibody) versus mock-transfected parental cells; other panels compare calcein-1 loaded, purified leukocyte populations adhering to 293-CX₃C cells (diamonds) versus mock-transfected 293-pcDNA cells (squares).

chemotactic activity for T cells is likely to be contained in the chemokine domain. Two non-natural variants, a truncated CX₃C chemokine (the product of the 0.7-kb cDNA and containing only half of the mucin-like stalk) and a synthetic chemokine-only domain protein, showed similar specificity and potency for T-cell migration (see Supplementary Information), although these forms had reduced monocyte attractant activity (results not shown).

To test the function of the cell-surface-bound CX₃C chemokine, we first used flow-assisted cell sorting (FACS) analysis of stably transfected HEK293 cells to demonstrate expression of cell-surface CX₃C chemokine (Fig. 4b, top left). When monolayers of CX₃C

chemokine-expressing HEK293 cells (293-CX₃C) were incubated with purified populations of fluorescently labelled monocytes, T cells and neutrophils, membrane-bound CX₃C chemokine supported a time-dependent adhesion of monocytes and T cells but not neutrophils (Fig. 4b). HEK293 cells that had been selected identically, but not bearing CX₃C chemokine on their cell surface (293-pcDNA), did not promote adhesion. Adhesion was a function of the chemokine portion of the molecule because soluble synthetic protein representing the chemokine domain significantly diminished the effect. Thus, cell-surface CX₃C chemokine can, in the absence of other cellular stimuli, induce the adhesion of monocytes

and T cells to heterologous cells expressing membrane-bound CX₃C chemokine.

As closer inspection of one branch of the chemokine superfamily, the C branch, yielded evidence of another, the CX₃C, rather as in fractal geometry (where dynamic, self-referential branch patterns in chaotic systems are revealed at increasing scales of magnification), we have dubbed this new CX₃C chemokine 'fractalkine'. CX₃C may regulate adhesion and migration processes at the most important interface in leukocyte trafficking: the luminal surface of the endothelium lining circulatory vessels and the surface of leukocytes moving through those vessels. The recruitment of leukocytes from blood to tissues during normal recirculation and inflammatory processes has been divided into four steps: (1) primary adhesion and rolling; (2) activation and arrest; (3) firm adhesion; and (4) diapedesis. These four steps are thought to be mediated through the concerted actions of cellular adhesion molecules and chemokines^{16–19}. Chemokines are thought to be involved in both rapid activation of the rolling leukocyte and in its locomotion into the tissue spaces. This model presupposes that chemokines are sequestered and immobilized on the endothelial cell surface, perhaps by binding to a variety of glycosaminoglycan moieties and proteoglycans^{19–21}. The structure of fractalkine and its inducible expression on endothelial cells suggest that it could function on endothelium by being 'presented' on its own extended stalk, obviating the need for other components of the extracellular matrix. This could promote the transient activation and adhesion of monocytes and T cells to the endothelium, facilitating step (2) of the trafficking process, the transition from leukocyte rolling to firm adhesion. Moreover, cleavage of CX₃C chemokine from the membrane might function as a regulatory mechanism to modulate trafficking at the local site. □

Methods

Sequence analysis. BLAST searches⁸ of the NCBI sequence databases with lymphotactin⁹ retrieved a set of related CX₃C chemokine ESTs from human (Genbank accession numbers R86674, Z44443, F06736, H06104, R87691, T31247 and H82247) and mouse (R75309) brain cDNA libraries. An EST clone, H06104, generated by the IMAGE consortium¹⁰ was obtained from Research Genetics (Huntsville, Alabama) as a *NotI*–*HindIII* insert in the *Lafmid* BA vector, and subjected to automated sequencing (ABI). A splice variant of CX₃C chemokine that results in an extended 3' UTR sequence was indicated by the overlapping sequence of a 1,998-bp human breast EST-derived (H14940) cDNA clone. Two closely related mouse brain-derived ESTs (R75038, W16003) probably mark the 3' end of a long CX₃C chemokine mRNA in mouse.

RNA expression, chromosomal localization and endothelial cell activation. Tissue northern blots were purchased from Clontech and Stratagene. Libraries for the cDNA southern blots were kindly provided by T. McClanahan and the novel cloning group at DNAX. All blots were probed with the 1.7-kb CX₃C chemokine cDNA comprising the entire coding domain, using high-stringency hybridizations. For endothelial cell northern blots, HUVEC and HPAEC were obtained from Clontech as frozen samples and seeded into T175 culture flasks at passage 2. At confluence, cells were incubated for 24 h with human recombinant TNF- α , (10 ng ml⁻¹; Boehringer Mannheim), and removed from culture flasks using 0.5 mM EDTA. mRNA was purified from 10⁷ cells using a Fast-Track kit (Invitrogen). For initial chromosomal localization, PCR was done on genomic DNAs (250 ng) from the BIPS human X mouse somatic cell-hybrid panel. The Clontech Advantage PCR kit was used with the PCR primers: ATGGACGAGTCTGTGGTCTGGAG (sense) and AGGCAATCGGAAAAGGTCCG (antisense). A 'touchdown' program was used with 10 cycles of annealing at 65 °C for 2 min and elongation at 72 °C for 2 min, followed by an additional 25 cycles in which the annealing temperature was 55 °C for 2 min 30 s. PCR products were analysed by electrophoresis on 1% agarose gels, revealing the presence of a diagnostic 207-bp band. The localization to chromosome 16 was confirmed by analysing three independently derived somatic-cell hybrids containing human chromosome 16: som.hyb.1, a SM106 murine A9 cell $\times$ human fibroblast line carrying human chromosome

16 (obtained from BIPS); som.hyb.2, a house MEL somatic-cell hybrid JYS-4 monochromosomal for human 16; som.hyb.3, a mouse MEL somatic-cell hybrid JLT585-P3 monochromosomal for human 16 (kindly provided by D. Higgs).

Biochemistry. Transfection, biosynthetic labelling, supernatant collection and western blot analysis have all been described²². The expression vector used for all constructs was pcDNA 3 (Invitrogen). Rabbit polyclonal antiserum was raised against the synthetic peptide CADPKEQWVKDAMQHLDRQAAL, coupled at the N terminus to KLH, at a 1:500 dilution²². Preimmune serum from the same animal was used as a negative control. IgG was purified in both cases by standard methods.

Chemotaxis and adhesion. Human T lymphocytes were purified (>98%) by Ficoll density gradient centrifugation of buffy coats, and removal of monocytes and B cells by negative selection with magnetic-bead-coupled antibodies. PMN leukocytes (>80% neutrophils) were prepared from freshly drawn blood sedimented on dextran (6% dextran 70 in 0.9% NaCl) to remove PBMC and hypotonic lysis of RBC. Monocytes were isolated from buffy coats following Ficoll separation of PBMC and removal of T and B lymphocytes by negative selection. All leukocyte populations were resuspended in serum-free 1640 medium and assayed immediately. Chemotaxis of leukocytes was done as described²². For T lymphocytes, 80- μ m-pore-size PVDF polycarbonate filters were used; for monocyte and PMN migration, 5- μ m and 3- μ m filters, respectively, were used. Adhesion assays were done on purified peripheral monocytes (isolated by negative selection using anti-CD2, CD3, CD8, CD19, CD20, CVD56 or CD67 T cells (purified on T-cell-enrichment columns; R&D Systems), or neutrophils (purified as above with and without erythrocyte lysis). Some assays were done using soluble synthetic protein comprising the chemokine domain of fractalkine (amino acids 1–76) and including the M2 flag epitope DYKDDDDK at the C terminus. This protein was custom-synthesized by Gryphon Sciences using native chemical ligation²³. Monolayers of 293 cells bearing the CX₃C chemokine (as well as mock-transfected, identically selected control 293 cells) growing in 96-well plates were overlaid with the various leukocyte populations that had been loaded with Calcein IAM fluorescent dye (Molecular Probes; 10 μ M final concentration). Plates were spun briefly, incubations proceeded for the times indicated, non-adherent cells were washed away, and the fluorescence intensity of each plate was obtained as a ratio of emission after excitation at 485 and 530 nm using a Cytofluor 2350 fluorimeter (Millipore). Each point represents a quadruplicate determination; assays are typical of $n = 10$ for each cell type.

Received 27 August 1996; accepted 17 January 1997.

1. Strieter, R. M. *et al.* *J. Immunol.* **156**, 3583–3586 (1996).
2. Baggiolini, M. & Dahinden, C. A. *Immunol. Today* **104**, 27–29 (1994).
3. Schall, T. J. & Bacon, K. B. *Curr. Opin. Immunol.* **6**, 865–873 (1994).
4. Weiss, R. A. & Clapham, P. R. *Nature* **381**, 647–648 (1996).
5. Pemack, B. P. & Schall, T. J. *Nature Med.* **2**, 1174–1178 (1996).
6. Clore, G. M. & Gronenborn, A. M. *FASEB J.* **9**, 57–62 (1995).
7. Wells, T. N. *et al.* *J. Leuk. Biol.* **59**, 53–60 (1996).
8. Altschul, S. F., Boguski, M. S., Gish, W. & Wootton, J. C. *Nature Genet.* **6**, 119–129 (1994).
9. Kelner, G. S. *et al.* *Science* **266**, 1395–1399 (1994).
10. Lennon, G., Auffray, C., Polymeropoulos, M. & Soares, M. B. *Genomics* **33**, 151–152 (1996).
11. Gendler, S. J. & Spicer, A. P. *Annu. Rev. Physiol.* **57**, 607–634 (1995).
12. Hansen, J. E. *et al.* *Biochem. J.* **308**, 601–613 (1995).
13. Shimizu, Y. & Shaw, S. *Nature* **366**, 630–631 (1993).
14. Bernfield, M. *et al.* *Annu. Rev. Cell Biol.* **8**, 365–393 (1992).
15. Vyas, P., Vickers, M. A., Picketts, D. J. & Higgs, D. R. *Genomics* **29**, 679–689 (1995).
16. Springer, T. A. *Cell* **76**, 301–314 (1994).
17. Lasky, L. A. *Science* **258**, 964–969 (1992).
18. Butcher, E. C. & Picker, L. J. *Science* **272**, 60–66 (1996).
19. Webb, L. M. C., Ehrengreuter, M. U., Clark-Lewis, I., Baggiolini, M. & Rot, A. *Proc. Natl. Acad. Sci. USA* **90**, 7158–7162 (1993).
20. Tanaka, Y., Adams, D. H. & Shaw, S. *Immunol. Today* **361**, 79–82 (1993).
21. Witt, D. P. & Lander, A. D. *Curr. Biol.* **4**, 394–400 (1994).
22. Schall, T. J., Bacon, K., Toy, K. J. & Goeddel, D. V. *Nature* **347**, 669–671 (1990).
23. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. *Science* **266**, 776–779 (1994).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank D. Liggett for oligonucleotide synthesis; C. Huffine, A. Helms and D. Gorman for automated DNA sequencing; T. McClanahan, K. B. Bacon, J. Mattson and M. Sterling for cDNA libraries; G. Burgett for help in preparing the figures; R. Kastelein, F. Rock, T. Handel and L. Mizoue for helpful discussions; and D. Higgs for cell lines. The DNAX Research Institute is supported by Schering-Plough.

Correspondence and requests for materials to T.S. (e-mail: schall@dnax.org). Sequence has been deposited with Genbank, accession number U84487.

CCR3 and CCR5 are co-receptors for HIV-1 infection of microglia

Jianglin He^{*†}, Youzhi Chen^{*}, Michael Farzan^{*}, Hyeryun Choe^{*†}, Asa Ohagen^{*†}, Suzanne Gartner[‡], Jorge Busciglio^{§||}, Xiaoyu Yang^{*†}, Wolfgang Hofmann^{*†}, Walter Newman[†], Charles R. Mackay[†], Joseph Sodroski^{*†} & Dana Gabuzda^{*§}

^{*} Division of Human Retrovirology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA

[†] Departments of Pathology and § Neurology, Harvard Medical School, Boston, Massachusetts 02115, USA

[‡] Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

^{§||} Department of Neurology, Children's Hospital Medical Center, Boston, Massachusetts 02115, USA

[¶] LeukoSite, Inc., Cambridge, Massachusetts 02142, USA

Several members of the chemokine receptor family are used together with CD4 for HIV-1 entry into target cells^{1–6}. T cell line-tropic (T-tropic) HIV-1 viruses use the chemokine receptor CXCR4 as a co-receptor¹, whereas macrophage-tropic (M-tropic) primary viruses use CCR5 (refs 2–6). Individuals with defective CCR5 alleles exhibit resistance to HIV-1 infection^{7,8}, suggesting that CCR5 has an important role *in vivo* in HIV-1 replication. A subset of primary viruses can use CCR3 as well as CCR5 as a co-receptor^{5,6}, but the *in vivo* contribution of CCR3 to HIV-1 infection and pathogenesis is unknown. HIV-1 infects the central nervous system (CNS) and causes the dementia associated with AIDS⁹. Here we report that the major target cells for HIV-1 infection in the CNS, the microglia^{9–11}, express both CCR3 and CCR5. The CCR3 ligand, eotaxin, and an anti-CCR3 antibody inhibited HIV-1 infection of microglia, as did MIP-1β, which is a CCR5 ligand. Our results suggest that both CCR3 and CCR5 promote efficient infection of the CNS by HIV-1.

The chemokine receptors that mediate HIV-1 entry in cultured cells vary in their tissue distribution. CCR5 is expressed in primary monocyte/macrophages, primary T cells, and granulocyte precursors²⁴. CXCR4 is expressed in a broader range of tissues and cell types, including the brain and T cells¹. The expression of CCR3, however, has so far only been detected in eosinophils^{12–14}. To study the expression of CCR5, CXCR4 and CCR3 in the brain, we performed reverse transcription–polymerase chain reaction (RT–PCR) analysis of primary human brain cultures using oligonucleotide primer pairs specific for each chemokine receptor. CCR5, CXCR4 and CCR3 transcripts were detected in primary brain cultures (Fig. 1a). The identity of these transcripts was confirmed by DNA sequencing of the PCR products (data not shown). CCR3 transcripts were relatively abundant. We therefore examined whether CCR3 is expressed in microglia by double immunofluorescence staining with an anti-CCR3 monoclonal antibody¹⁴ and the microglial cell marker *Ricinus communis* agglutinin I (RCA-1). CCR3 was expressed in >95% of microglia in primary brain cultures, but was not detected in other brain cell types (Fig. 1b). CCR3 expression in microglia was also detected in autopsy brain tissue from two normal adults and two patients with Alzheimer's disease (Fig. 1c and data not shown). Similar experiments using an anti-CCR5 monoclonal antibody showed that microglia also express CCR5 (data not shown). The microglia in these specimens did not stain with monoclonal antibodies directed against galactosyl ceramide or glial fibrillary acidic protein (data not shown).

Most of the HIV-1-infected cells in the brain are microglia^{9–11}.

Astrocytes are infected at a very low level^{11,15,16}. HIV-1 entry into microglia is CD4 dependent, but entry into astrocytes and some neurally derived cell lines is CD4 independent^{15,16}. HIV-1 viruses that infect the CNS are M-tropic^{10,17–20}. Particular Env sequences may be associated with brain infection or dementia^{17,18}, although specific determinants of HIV-1 neurotropism have not been identified. We examined the ability of HIV-1 Env proteins with different chemokine receptor specificity to mediate virus entry into primary human brain cultures containing microglia and astrocytes. We used ADA (ref. 21), YU2 (ref. 22) and Br20-4 (ref. 5) as M-tropic Env proteins, and HXB2 as a T-tropic Env. We also tested the dual-tropic 89.6 Env, which supports broader HIV-1 tropism²³. The chemokine receptor use and biological properties of these HIV-1 Envs are summarized in Table 1. An *env*-defective HIV-1 reporter virus that expresses the green fluorescent protein (GFP) was constructed. Complementation of this virus with HIV-1 Env allows a single round of infection to occur, and infected cells can be detected by expression of GFP. Recombinant HIV-1 GFP reporter viruses were used to infect primary human fetal brain cultures. To identify HIV-1-infected cells *in situ*, GFP fluorescence was visualized in combination with cell-specific markers for microglia and astrocytes (Fig. 2); the results are summarized in Table 1. The most efficient infection of microglia was mediated by the YU2, ADA and 89.6 Env proteins followed by Br20-4 (Table 1). Microglial infection mediated by the HXB2 Env was relatively inefficient. A very small fraction of astrocytes (2–4%) was infected by viruses with the YU2, ADA, 89.6, Br20-4 or HXB2 Env proteins, indicating that a low efficiency of HIV-1 entry into astrocytes can be mediated by either M-tropic or T-tropic Env proteins. In contrast, virus with the amphotropic murine leukaemia virus (A-MLV) Env infected 10–30% of the astrocyte population (data not shown). The level of GFP expression in astrocytes was generally lower than that seen in microglia, consistent with previous studies which demonstrated that HIV-1 gene expression in astrocytes is restricted^{11,16}. These results demonstrate that M-tropic HIV-1 viruses that use CCR5 and CCR3 as co-receptors infect microglia efficiently, whereas a T-tropic virus that uses CXCR4 is relatively inefficient.

The natural ligands for the chemokine receptors have been shown to inhibit the entry of particular HIV-1 isolates that use these receptors^{1–6,24–26}. MIP-1α, MIP-1β and RANTES, which are chemokines that bind CCR5, inhibit infection by M-tropic but not T-tropic HIV-1 isolates^{2–4,24}. Similarly, stromal cell-derived factor-1 (SDF-1) inhibits entry by T-tropic or dual-tropic HIV-1 isolates that use CXCR4 (refs 25, 26). Infection by the subset of M-tropic isolates that use CCR3 as a co-receptor is inhibited by eotaxin⁵, the main CCR3 ligand^{12–14}. We tested the ability of chemokines to inhibit HIV-1 infection of primary brain cultures. Chemokine blocking experiments were performed using *env*-defective HIV-1 luciferase reporter viruses pseudotyped with different HIV-1 Env proteins; the A-MLV Env was used to control for possible nonspecific inhibitory effects. The efficiency of the early phase of virus replication during a single-cycle infection was determined by measuring luciferase activity in the primary brain cultures 72 h after infection (Fig. 3). The relative efficiency of virus entry mediated by the different HIV-1 Env proteins was similar to that observed using the HIV-1 GFP reporter viruses. We initially tested MIP-1β, eotaxin and SDF-1, as these chemokines show selective binding to CCR5, CCR3 or CXCR4, respectively. MIP-1β and eotaxin inhibited infection with HIV-1 luciferase reporter viruses containing the YU2 or ADA Envs by 70–80% (Fig. 3a). In contrast, infection with viruses containing the Br20-4 or 89.6 Env was inhibited by MIP-1β, but was not significantly inhibited by eotaxin. MIP-1β and eotaxin had no inhibitory effect on the virus with the HXB2 Env. SDF-1 inhibited infection by viruses with the 89.6 or HXB2 Env, but not the YU2, ADA or Br20-4 Env. MCP-1, which does not bind CCR5, CCR3 or CXCR4, had no inhibitory effect on any of the viruses tested. None of the chemokines tested inhibited infection by virus with the A-MLV

Env (Fig. 3b). These results are consistent with the known chemokine receptor use of the different HIV-1 Env proteins (Table 1), and suggest that chemokines that bind either CCR5, CCR3 or CXCR4 can inhibit infection of CNS target cells by HIV-1 isolates with the same chemokine receptor selectivity.

Our finding that eotaxin showed potent blocking activity for viruses bearing the YU2 or ADA Env proteins led us to examine further the possibility that CCR3 is a co-receptor for HIV-1 infection of microglia. To address this question, we tested the blocking activity of the 7B11 monoclonal antibody, which specifically recognizes CCR3 and completely blocks the binding and signalling of the known CCR3 ligands¹⁴. To test whether the 7B11 antibody inhibits HIV-1 entry into CCR3-expressing cells and to confirm its specificity, Cf2Th canine thymocytes transfected with

plasmids expressing CD4 and either CCR3 or CCR5 (ref. 5) were infected with recombinant HIV-1 chloramphenicol acetyltransferase (CAT) reporter viruses containing the YU2 Env in the absence or presence of the antibody. Cf2Th cells transfected with plasmids expressing CD4 and CXCR4 were infected with the HIV-1 CAT reporter virus containing the HXB2 Env in parallel experiments. The 7B11 monoclonal antibody blocked infection of cells expressing CCR3, but not those expressing CCR5 or CXCR4, as determined by inhibition of CAT activity in the target cells 60 h after infection (Fig. 3c). These findings indicate that the anti-HIV-1 effect of the 7B11 antibody is specific for infections in which CCR3 is used as a co-receptor.

We then examined the blocking activity of the anti-CCR3 monoclonal antibody in primary brain cultures. The anti-CCR3 antibody inhibited infection by HIV-1 luciferase reporter viruses

Figure 1 Expression of CCR3 in microglia. **a**, RT-PCR analysis of CXCR4, CCR3 and CCR5 transcripts in primary brain cultures. The 371-bp CXCR4, 313-bp CCR3, and 506-bp CCR5 PCR products are shown. The positions of molecular-weight markers (M) are indicated on the left. To control for possible genomic DNA contamination, reactions with and without reverse transcriptase (RT) were performed in parallel. To confirm that the assay conditions were within the linear range, the PCR amplification was compared with that of CXCR4, CCR3 and CCR5 expression plasmids using a 10-fold dilution series, starting at 1 ng of plasmid DNA. **b**, **c**, Double immunofluorescence staining of CCR3 in microglia. Primary brain cultures (**b**) or frozen human brain tissue from a patient with Alzheimer's disease (**c**) were fixed and stained with the 7B11 anti-CCR3 monoclonal antibody raised against CCR3 transfectants¹⁴ and a fluorescein-conjugated secondary antibody followed by staining with RCA-1 conjugated to rhodamine, a marker for microglia. Top, fluorescein-specific fluorescence; middle, rhodamine-specific fluorescence; bottom, double exposures in which the colocalization of fluorescein and rhodamine appears yellow. Original magnification, $\times 1,000$.

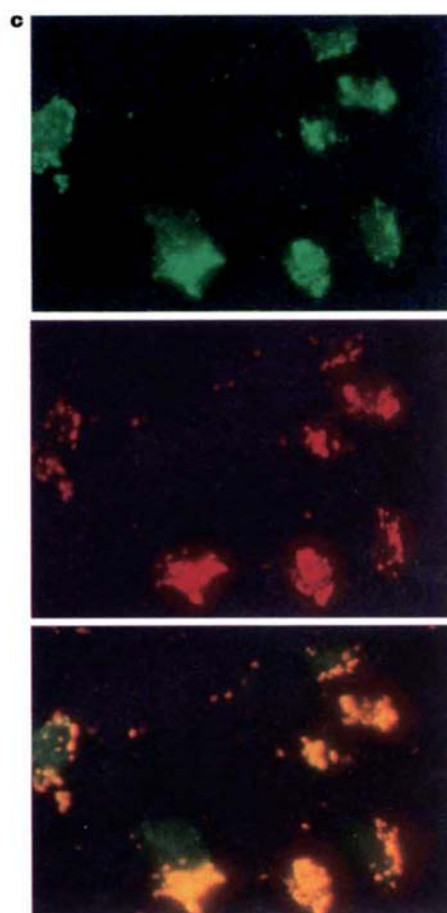
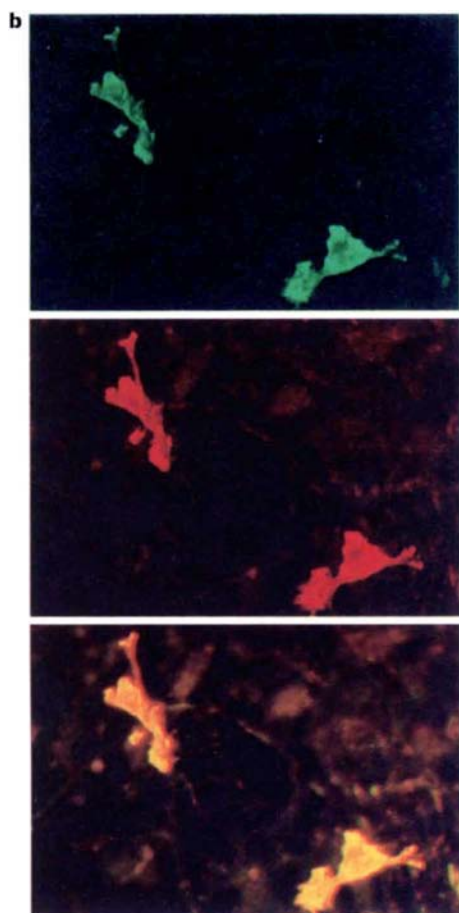
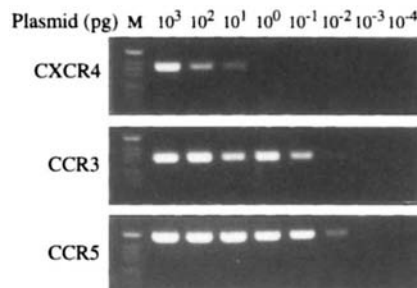
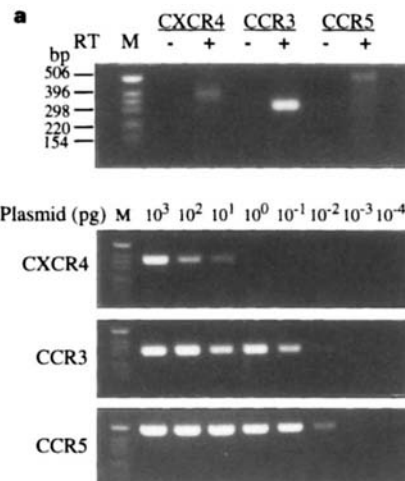


Table 1 HIV-1 infection of microglia and astrocytes

HIV-1 envelope	Classification	Chemokine receptor use	HIV-1 infection (%)	
			Microglia	Astrocytes
YU2	M-tropic primary, NSI	CCR5, CCR3	44 ± 4	3.8 ± 0.8
ADA	M-tropic primary, NSI	CCR5, CCR3	42 ± 6	3.3 ± 0.2
Br20-4	M-tropic primary, NSI	CCR5	36 ± 1	3.2 ± 0.8
89.6	Dual-tropic primary, SI	CXCR4, CCR5, (low CCR3)	48 ± 2	2.4 ± 0.9
HXB2	T-tropic laboratory adapted, SI	CXCR4	10 ± 0	3.0 ± 0.5

HIV-1 infection in primary brain cultures mediated by Env proteins with different chemokine receptor use. The efficiency of HIV-1 entry was determined by single-cycle infections using HIV-1-GFP reporter viruses pseudotyped with different HIV-1 Env proteins. The chemokine receptor use is derived from ref. 5. The efficiency of HIV-1 entry for each Env is expressed as the percentage of CD68⁺ (microglia) or GFAP⁺ (astrocytes) cells labelled with GFP, as determined by counting 10 random fields (magnification, × 200). Data shown are the mean ± s.d. of duplicates from a representative experiment. Results were comparable in 4 independent experiments, although the maximum HIV-1 entry ranged from 30 to 45% for microglia, and from 2 to 5% for astrocytes between different donors. Abbreviations: NSI, non-syncytium-inducing; SI, syncytium-inducing.

with the YU2 or ADA Env with an efficiency comparable to that of eotaxin (Fig. 3c). To test whether ligands that bind to both CCR3 and CCR5 show greater blocking activity than ligands selective for either CCR3 or CCR5, we also tested the blocking activity of RANTES, a ligand for both receptors, and the combination of eotaxin and MIP-1 β . The blocking activity of RANTES alone, or the combination of eotaxin and MIP-1 β , was not significantly greater than that of eotaxin alone (Fig. 3c). The chemokines and the 7B11 antibody had no inhibitory effect on viruses with the A-MLV Env (Fig. 3b). We also tested the inhibitory effect of the anti-CD4 monoclonal antibody OKT4A, which blocks binding of CD4 to gp120. The OKT4A antibody inhibited infection by HIV-1 luciferase viruses containing the YU2, ADA, Br20-4 or 89.6 Env proteins by 70–80%, and infection by virus with the HXB2 Env was inhibited by 50% (data not shown). In contrast, the infection of the CD4⁺ Jurkat T cell line was inhibited by >95% under the same assay conditions.

These experiments demonstrated that CCR3 and CCR5 are used for HIV-1 infection of primary brain cultures. To confirm that CCR3 and CCR5 are used for HIV-1 infection of microglia, the inhibitory effects of eotaxin, MIP-1 β and the 7B11 anti-CCR3 antibody were examined by direct visualization of infected target cells using HIV-1 GFP reporter viruses. Double labelling with cell-specific markers showed that eotaxin, MIP-1 β and anti-CCR3 inhibited 70–80% of microglial infection by HIV-1 GFP reporter viruses with the YU2 Env (Fig. 3c), which was cloned directly from brain²². To examine use of CCR3 and CCR5 by other primary neurotropic HIV-1 isolates, we tested the inhibitory effects of eotaxin, MIP-1 β and RANTES on infection by HIV-1 viruses isolated directly from the brain tissue of patients with AIDS. JR-FL and DS-br were isolated from two adult AIDS patients with dementia, and KJ-br was isolated from a paediatric AIDS patient with encephalopathy^{20,27}. Eotaxin, MIP-1 β and RANTES blocked infection of primary brain cultures by JR-FL and KJ-br by 50–75%, whereas infection by DS-Br was inhibited by 30–50% (Fig. 3d). To test directly whether CCR3 is used for HIV-1 entry mediated by the JR-FL Env, we examined the ability of HIV-1 CAT reporter viruses with the JR-FL Env to infect Cf2Th cells transfected with plasmids expressing CD4 and either CCR3, CCR4 or CCR5 (ref. 5). Coexpression of either CCR3 or CCR5 significantly enhanced infection mediated by the JR-FL Env, similar to their effects on infection mediated by the YU2 Env (Fig. 3e). These results suggest that CCR3 and CCR5 have an important role in the infection of microglia by HIV-1.

Our results suggest that both CCR3 and CCR5 serve as major co-receptors with CD4 for HIV-1 infection of brain microglia. The M-tropic primary HIV-1 isolates infect peripheral blood monocyte/macrophages and brain-derived microglia more efficiently than do T-tropic isolates. The use of both CCR3 and CCR5 for microglial entry differs from the infection of blood-derived monocyte/macrophages, which express CCR5 (refs 2 and 4) but do not express CCR3. Our results provide an explanation for the observation that subsets of M-tropic primary HIV-1 isolates use CCR3 as well as CCR5 as a second receptor^{5,6}. Our study predicts that such viruses reside in the CNS, and that the unique pattern of chemokine receptor expression in microglia selects for particular features of the viral envelope glycoproteins. The finding that the YU2 virus, which was derived from brain²², shows more efficient CCR3 usage than M-tropic or dual-tropic isolates derived from the blood (ADA, Br20-4, 89.6 and ELI), and the observation that some primary isolates show preferential replication in microglia compared to blood-derived monocyte/macrophages¹⁹, are consistent with this prediction. The use of both CCR3 and CCR5 for HIV-1 entry into microglia is also supported by our finding that ligands for both receptors can inhibit infection by several primary neurotropic HIV-1 isolates. The finding that MIP-1 α and MIP-1 β are induced in the brain of AIDS patients²⁸ raises the possibility that viruses continue to replicate in the CNS in the presence of these CCR5 ligands by

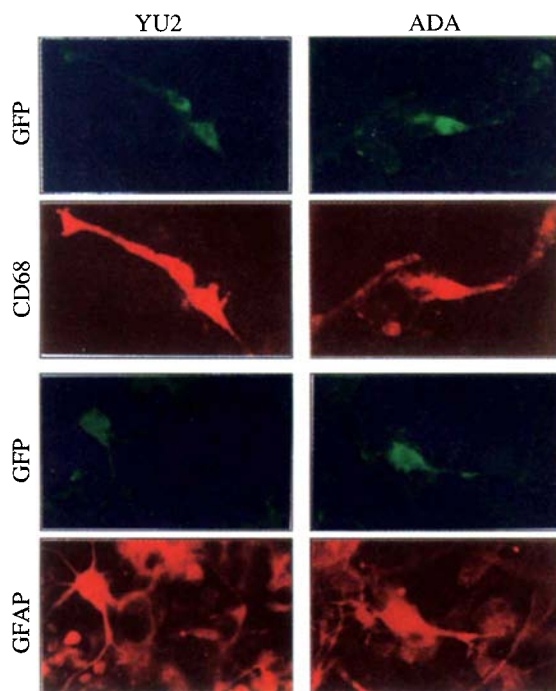


Figure 2 *In situ* detection of HIV-1 entry into microglia and astrocytes. Primary human fetal brain cultures were infected by HIV-1 GFP reporter viruses containing the HIV-1 YU2 or ADA Env proteins. The cultures were fixed 72 h after infection, stained with antibodies directed against the astrocyte marker GFAP or the microglial marker CD68 and rhodamine-conjugated secondary antibodies, and visualized with fluorescein or rhodamine wavelength filters to detect virus infection and expression of cell markers, respectively. Original magnification, × 400. Results were similar in four independent experiments.

using CCR3 for virus entry. In contrast to previous reports^{2,3}, our results indicate that JR-FL can use both CCR3 and CCR5 for virus entry. This discrepancy might reflect differences in the levels of CCR3 expression in the target cells used in these studies.

Our data indicate that the CCR3 and CCR5 ligands interfere with HIV-1 entry into microglia more than might be expected from the inhibition of independently functioning co-receptors expressed at saturating levels. The possibility of cooperativity between CCR3 and CCR5 in HIV-1 entry or between ligand-mediated mechanisms of HIV-1 inhibition merits further investigation. The identification of CCR3 as a co-receptor contributing to efficient HIV-1 infection of

microglia may suggest new therapeutic strategies to inhibit HIV-1 replication in the CNS.

Methods

Primary brain cultures. Primary human brain cultures were prepared from fetal abortuses at 13–18 weeks, plated in 24-well plates (250,000 cells per well), and maintained in DMEM containing 10% calf serum for 10–21 days before infection or detection of chemokine receptor expression²⁹. Tissue was procured in accordance with institutional regulations. These cultures contain a mixture of astrocytes (70–90%), neurons (10–30%), microglial cells (1–5%) and fibroblasts (1–5%)²⁹.

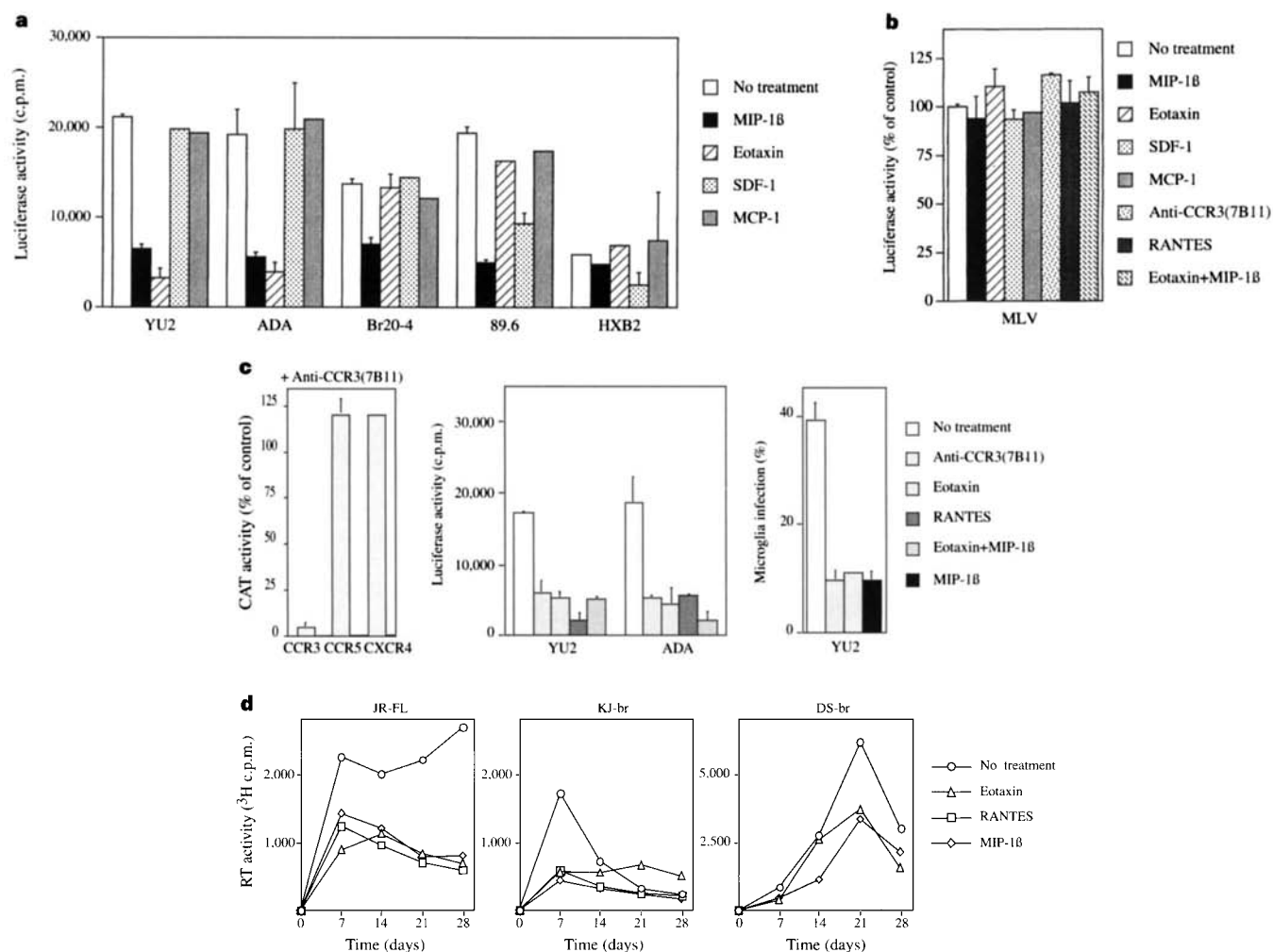
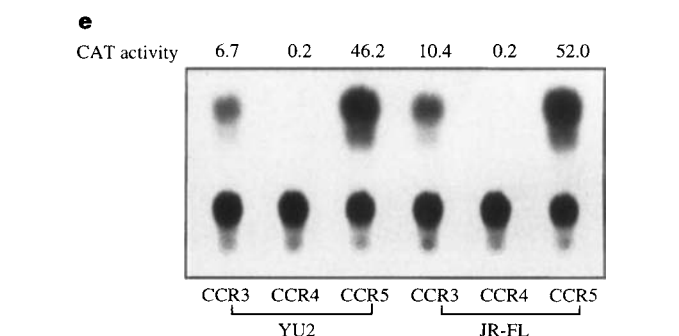


Figure 3 Chemokine and anti-CCR3 inhibition of HIV-1 infection of primary brain cultures. For **a–c** (middle and right), cultures were pre-incubated with the individual chemokines shown, eotaxin and MIP-1 β in combination, or the 7B11 anti-CCR3 monoclonal antibody for 1 h before infection with HIV-1 luciferase or GFP reporter viruses containing the YU2, ADA, BR20-4, 89.6, HXB2 or A-MLV Env proteins. Chemokines were used at 500 ng ml⁻¹ except for SDF-1, which was used at 2.5 μ g ml⁻¹. The 7B11 antibody was used at 20 μ g ml⁻¹. Luciferase activity in the cell lysates was measured 72 h after infection. **c** (right), The percentage of HIV-1-infected microglia was determined by counting the percentage of GFP⁺, CD68⁺ cells (as in Table 1). **c** (left), Cf2Th-CD4 canine thymocytes were transfected with plasmids expressing either CCR3, CCR5 or CXCR4 and infected with HIV-1 CAT reporter viruses containing the YU2 (CCR3 and CCR5) or HXB2 (CXCR4) Env in the absence or presence of the 7B11 antibody (3 μ g ml⁻¹). Results are indicated as the percentage of chloramphenicol conversion observed in the absence of antibody. Mean $\pm$ s.d. of duplicates are shown. Similar results were obtained in two or three independent experiments. **d**, Cultures were preincubated with the individual chemokines shown as in **a–c** before infection with the primary



brain HIV-1 isolates JR-FL, KJ-br or DS-br. **e**, Cf2Th-CD4 canine thymocytes were transfected with plasmids expressing either CCR3, CCR4 or CCR5 and infected by HIV-1 CAT reporter viruses with the YU2 or JR-FL Env. Above, the CAT activity in the cell lysates is expressed as the percentage of chloramphenicol conversion.

Production of HIV-1 viruses. Recombinant GFP or luciferase reporter viruses were generated by co-transfection of COS-7 or 293 cells with 20 µg of pNL4-3env⁻ GFP or pNL4-3env⁻ LUC and 4 µg of pSVIIIenv plasmids encoding different HIV-1 Env proteins or pSVMLVenv using the calcium phosphate method³⁰. The pNL4-3env⁻ GFP plasmid, which encodes full-length NL4-3 HIV-1 proviral DNA with a frameshift in *env* and expresses GFP in place of *nef*, was constructed by replacing the alkaline phosphatase (AP) gene in pHIV-AP³⁰ with the GFP gene using *NotI* and *XhoI* restriction sites introduced into pEGFP-1 (Clontech) by PCR amplification. pNL4-3env⁻ LUC, which encodes full-length *env*-defective NL4-3 HIV-1 proviral DNA and expresses the luciferase enzyme, was similarly constructed by replacement of the AP gene in pHIV-AP with the luciferase gene from pGEM-luc (Promega). The pSVIIIenv plasmids, which express Env proteins from different strains of HIV-1, and the pSVMLVenv plasmid, which expresses the amphotropic murine leukaemia virus envelope glycoproteins, have been described⁵. The JR-FL Env expression plasmid was constructed in pcDNA3 (Invitrogen). Supernatants containing virus were collected 48 h after transfection, quantified by RT assay²⁹, and stored at -80 °C. A similar method was used to produce HIV-1 CAT reporter viruses pseudotyped with different HIV-1 Env proteins⁵. JR-FL virus stock was prepared from the supernatants of infected PHA-stimulated peripheral blood mononuclear cells. KJ-br and DS-br virus stocks were prepared from the supernatants of infected peripheral blood-derived monocyte/macrophages²⁰. Cell-free culture supernatants were quantified by RT assay and stored at -70 °C.

HIV-1 entry and chemokine inhibition assays. Primary brain cultures were infected by incubation with recombinant virus (5,000 RT units) pseudotyped with the different Env proteins in 1 ml of medium. After 16 h incubation at 37 °C, the medium was removed and the cultures were washed three times before addition of fresh medium. After an additional 72 h at 37 °C, the cells were fixed in 4% paraformaldehyde in PBS for visualization of GFP and immunofluorescence staining or lysed for determination of luciferase activity using commercially available reagents (Promega) after normalization for the same protein concentration. For chemokine or antibody inhibition experiments, cultures were preincubated for 60 min at 37 °C in 0.5 ml medium with chemokine (500 ng ml⁻¹, except for SDF-1, which was used at 2.5 µg ml⁻¹) or antibody (20 µg ml⁻¹) before infection. Subsequently, infections were performed by addition of recombinant virus in an equal volume of medium without additional chemokine or antibody for 16 h at 37 °C. The cells were then washed and fresh medium without chemokine or antibody was added. The cells were cultured for an additional 72 h at 37 °C, and then processed as described above. Results were expressed as the luciferase activity (c.p.m.) after subtraction of the mean background levels obtained with no Env. A similar infection protocol was used to test chemokine inhibition of primary brain isolates, except that the virus inoculum was 20,000 RT units and the results were monitored by measured RT activity (³H c.p.m. ml⁻¹) in the culture supernatants over a 28-day time course. As in the above protocol, chemokine was added only during the initial period of infection. For antibody inhibition experiments, Cf2Th canine thymocytes were transfected with plasmids expressing CD4 and either CCR3, CCR5 or CXCR4 (ref. 5). The transfected cells were incubated either in the absence or presence of the 7B11 antibody (3 µg ml⁻¹) for 1 h at 37 °C. The cells subsequently were infected by HIV-1 CAT reporter viruses pseudotyped with an HIV-1 Env (YU2 Env for the CCR3 and CCR5 transfectants, HXB2 Env for the CXCR4 transfectants) and CAT activity was measured in the transfectants cells 60 h after infection⁵. A similar protocol was used to assess the efficiency of infection by HIV-1 CAT reporter viruses pseudotyped with the JR-FL or YU2 Env into Cf2Th cells transfected with plasmids expressing CD4 and either CCR3, CCR4 or CCR5 (ref. 5).

Immunofluorescence staining. Immunofluorescence staining of fixed cultures with the indicated primary antibodies followed by FITC- or rhodamine-conjugated secondary antibodies (Sigma) was performed²⁹. The dilutions for the primary antibodies were: mouse anti-CD68 monoclonal (EBM11) (Dako), 1:10; rabbit anti-GFAP (Sigma), 1:100; mouse anti-CCR3 monoclonal (7B11)¹⁴, 1:50. Labelling with RCA-1 conjugated to rhodamine (10 µg ml⁻¹) (Vector) was performed for 1 h at room temperature. For staining of primary brain cultures, the anti-CCR3 monoclonal antibody was detected by incubation with biotinylated horse anti-mouse IgG (5 µg ml⁻¹) for 30 min followed by avidin-FITC (20 µg ml⁻¹) for 15 min (Vector).

RT-PCR. Total RNA was isolated (Oncor) and 0.5 µg was used for cDNA synthesis using Superscript II RNase H reverse transcriptase and random hexamer primers (Gibco-BRL); one-tenth of this reaction was used as a template for PCR amplification with AmpliTaq DNA polymerase (Perkin-Elmer) (CXCR4 primers, 5'-GACCGCTACCTGGCCATT-3' and 5'-GTTGTAGGG-CAGCCA GCA-3'; CCR5 primers, 5'-AATCTTCTTCATCATCTCC-3' and 5'-TCTC TGTCACCTGCATAGC-3'; CCR3 primers, 5'-TCCTTCTCTTCTT-CCTATCA ATC-3' and 5'-GGCAATTTTCTGCATCTG-3') for 30 cycles at 94 °C for 1 min, 50 °C (for CXCR4) or 55 °C (for CCR5 and CCR3) for 1 min, and 72 °C for 1 min.

Received 27 November 1996; accepted 17 January 1997.

- Feng, Y., Broder, C. C., Kennedy, P. E. & Berger, E. A. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* **272**, 872-877 (1996).
- Deng, H. K. et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**, 661-666 (1996).
- Dragic, T. et al. HIV-1 entry into CD4⁺ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667-673 (1996).
- Alkhatib, G. et al. CC CKR5: A RANTES, MIP-1α, MIP-1β receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* **272**, 1955-1958 (1996).
- Choe, H. et al. The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* **85**, 1135-1148 (1996).
- Doranz, B. J. et al. A dual-tropic primary HIV-1 isolate that uses fusin and the β-chemokine receptors CCR-5, CCR-3, and CCR-2b as fusion cofactors. *Cell* **85**, 1149-1158 (1996).
- Liu, R. et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**, 367-377 (1996).
- Samson, M. et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* **382**, 722-725 (1996).
- Price, R. W. The brain in AIDS: central nervous system HIV-1 infection and AIDS dementia complex. *Science* **239**, 586-591 (1988).
- Watkins, B. A. et al. Specific tropism of HIV-1 for microglial cells in primary human brain cultures. *Science* **249**, 549-553 (1990).
- Takahashi, K. et al. *Ann. Neurol.* **39**, 705-711 (1996).
- Ponath, P. et al. Molecular cloning and characterization of a human eotaxin receptor expressed selectively on eosinophils. *J. Exp. Med.* **183**, 2437-2448 (1996).
- Daugherty, B. L. et al. Cloning, expression, and characterization of the human eosinophil eotaxin receptor. *J. Exp. Med.* **183**, 2349-2354 (1996).
- Heath, H. et al. Chemokine receptor usage by human eosinophils. The importance of CCR3 demonstrated using an antagonistic monoclonal antibody. *J. Clin. Invest.* (in press).
- Harouse, J. M. et al. CD4-independent infection of human neural cells by human immunodeficiency virus type 1. *J. Virol.* **63**, 2527-2533 (1989).
- Tornatore, C., Nath, A., Amemiya, K. & Major, E. O. Persistent human immunodeficiency virus type 1 infection in human fetal glial cells reactivated by T-cell factor(s) or by the cytokines tumor necrosis factor alpha and interleukin-1 beta. *J. Virol.* **65**, 6094-6100 (1991).
- Korber, B. T. M. et al. Genetic differences between blood- and brain-derived viral sequences from human immunodeficiency virus type 1-infected patients: evidence of conserved elements in the V3 region of the envelope protein of brain-derived sequences. *J. Virol.* **68**, 7467-7481 (1994).
- Power, C. et al. Demented and nondemented patients with AIDS differ in brain derived human immunodeficiency virus type 1 envelope sequences. *J. Virol.* **68**, 4643-4649 (1994).
- Strizki, J. M. et al. Infection of primary human microglia and monocyte-derived macrophages with human immunodeficiency virus type 1 isolates: evidence of differential tropism. *J. Virol.* **70**, 7654-7662 (1996).
- Gartner, S. & Popovic, M. Macrophage tropism of HIV-1. *AIDS Res. Hum. Retroviruses* **6**, 1017-1021 (1990).
- Westervelt, P., Gendelman, H. E. & Ratner, L. Identification of a determinant within the human immunodeficiency virus surface envelope glycoprotein critical for productive infection of primary monocytes. *Proc. Natl Acad. Sci. USA* **88**, 3097-3101 (1991).
- Li, Y. et al. Complete nucleotide sequence genome organization and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation. *J. Virol.* **66**, 6587-6600 (1992).
- Collman, R. et al. An infectious molecular clone of an unusual macrophage-tropic and highly cytopathic strain of human immunodeficiency virus type 1. *J. Virol.* **66**, 7517-7521 (1992).
- Cocchi, F. et al. Identification of RANTES, MIP-1α, and MIP-1β as the major HIV-suppressive factors produced by CD8⁺ T cells. *Science* **270**, 1811-1815 (1995).
- Bluel, C. C. et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature* **382**, 829-833 (1996).
- Oberlin, E. et al. The CXCR4 chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* **382**, 833-835 (1996).
- O'Brien, W. A. et al. HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside of the CD4-binding domain. *Nature* **348**, 69-73 (1990).
- Schmidt-Mayerson, H. et al. Human immunodeficiency virus type 1 infection alters chemokine β peptide expression in human monocytes: implications for recruitment of leukocytes into brain and lymph nodes. *Proc. Natl Acad. Sci. USA* **93**, 700-704 (1996).
- Shi, B. et al. Apoptosis induced by HIV-1 infection of the central nervous system. *J. Clin. Invest.* **98**, 1979-1990 (1996).
- He, J. & Landau, N. R. Use of a novel human immunodeficiency virus type 1 reporter virus expressing human placental alkaline phosphatase to detect an alternative viral receptor. *J. Virol.* **69**, 4587-4592 (1995).

Acknowledgements. We thank B. Yankner for discussion and primary brain cultures; I. Clark-Lewis for SDF-1; T. Springer, G. Karlsson and N. Sullivan for discussion; N. Landau for pHIV-AP plasmid; D. Littman for A-MLV env plasmid; B. Hahn, L. Ratner and R. Collman for HIV-1 envs; W. O'Brien for JR-FL env; and the NIH AIDS Research and Reference Reagent Program for JR-FL isolate. This work was supported by grants from the Pediatric AIDS Foundation (D.G.), NIH (D.G., I.S. and S.G.), and Swedish Medical Research Council (A.O.), and gifts from the G. Harold and Leila Mathers Charitable Foundation, the late William McCarty-Cooper, and the Friends 10. J.H. was supported in part by an NIH AIDS training grant.

Correspondence and requests for materials should be addressed to D.G. (e-mail: dana_gabuzda@dfci.harvard.edu). Requests for the 7B11 anti-CCR3 antibody should be directed to C.R.M. (e-mail: charles_mackay@leukosite.com).

Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement

Ismail Moarefi[†]§, Michelle LaFevre-Bernt[‡]§, Frank Sicheri^{*}, Morgan Huse^{*}, Chi-Hon Lee^{*}, John Kuriyan[†] & W. Todd Miller[‡]

^{*} Laboratories of Molecular Biophysics, [†] Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

[‡] Department of Physiology and Biophysics, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

§ These authors contributed equally to this work.

The protein Hck is a member of the Src family of non-receptor tyrosine kinases which is preferentially expressed in haematopoietic cells of the myeloid and B-lymphoid lineages^{1,2}. Src kinases are inhibited by tyrosine-phosphorylation at a carboxy-terminal site³⁻⁹. The SH2 domains of these enzymes play an essential role in this regulation by binding to the tyrosine-phosphorylated tail⁸⁻¹¹. The crystal structure of the downregulated form of Hck has been determined¹² and reveals that the SH2 domain regulates enzymatic activity indirectly; intramolecular interactions between the SH3 and catalytic domains appear to stabilize an inactive form of the kinase. Here we compare the roles of the SH2 and SH3 domains in modulating the activity of Hck in an investigation of the C-terminally phosphorylated form of the enzyme. We show that addition of the HIV-1 Nef protein, which is a high-affinity ligand for the Hck SH3 domain, to either the downregulated or activated form of Hck causes a large increase in Hck catalytic activity. The intact SH3-binding motif in Nef is crucial for Hck activation. Our results indicate that binding of the Hck SH3 domain by Nef causes a more marked activation of the enzyme than does binding of the SH2 domain, suggesting a new mechanism for regulation of the activity of tyrosine kinases.

Hck was coexpressed with C-terminal Src kinase (Csk) in Sf9 cells to produce Hck phosphorylated at Tyr 527 (c-Src numbering)¹² on the C terminus (the downregulated form⁸). Incubation of C-terminally phosphorylated Hck (10 μ M) with ATP increased the initial rate of phosphate transfer to a peptide substrate from a basal rate of 4.9 pmol min⁻¹ to a maximum of 54.2 pmol min⁻¹. The $t_{1/2}$ for activation under these conditions was $\sim$ 5 min (data not shown). When enzyme activity was assayed without ATP pretreatment, there was a lag in peptide phosphorylation (Fig. 1a), consistent with enzyme activation by autophosphorylation. This induction time was dependent on the enzyme concentration (Fig. 1a, inset). Using mass spectrometry, we confirmed that the site phosphorylated under these conditions is the major autophosphorylation site (Tyr 416 in Src). The time course of Hck autophosphorylation also showed a lag that corresponded to the induction time for peptide phosphorylation (Fig. 1b). These results suggest that, under our conditions, autophosphorylation of Hck is an intermolecular event. A similar conclusion has been obtained for Src in experiments with kinase-inactive mutants¹³ and by kinetic studies in a continuous assay¹⁴. In the latter system, the level of autophosphorylation at Src Tyr 416 was directly correlated with enzyme activation¹⁵. Using a form of Hck that had been dephosphorylated at Tyr 527, which had a specific activity 2.5 times higher than that of Hck pretreated with ATP, we found that the induction period for peptide phosphorylation was significantly reduced, and at higher concentrations ($\geq$ 0.5 μ M) the lag was absent altogether (Fig. 1c). The lag in autophosphorylation of dephosphorylated Hck was also significantly shortened relative to C-terminally phosphorylated Hck (data not shown). These results show that, as for Src¹⁵, the

lag in enzyme activation and autophosphorylation is dependent on C-terminal phosphorylation, consistent with the model that activation is due at least in part to release of the C-terminal phosphotyrosine from the SH2 domain of Hck.

The structure of downregulated Hck¹² indicates that the SH3 domain is important for stabilizing the inactive form of the enzyme. We tested whether displacement of the SH3 domain by another ligand would affect the enzyme activity. Nef, a 25K/27K protein

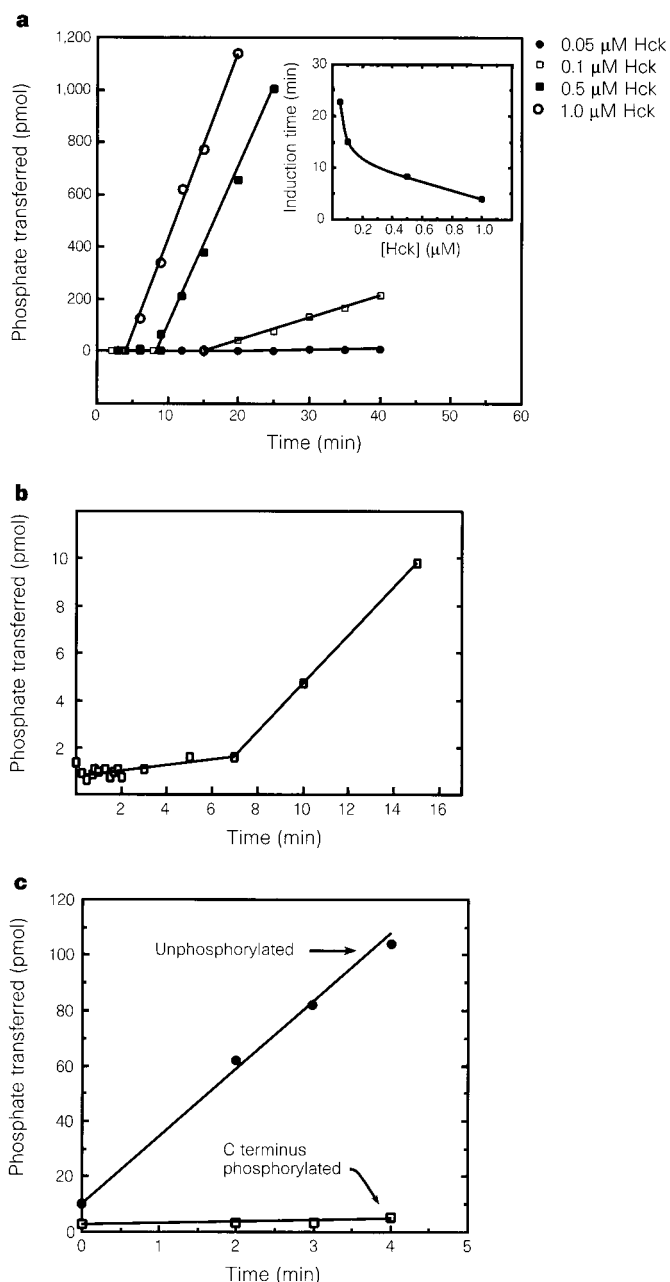


Figure 1 Activation of Hck. **a**, Activation of C-terminally phosphorylated Hck. Hck (no preincubation) was added to a kinase reaction containing peptide substrate. Aliquots were removed at the indicated times and ³²P incorporated into peptide was counted. Induction times were estimated from the intersection of the two phases of the progress curves of the reactions and are replotted in the inset. **b**, Lag in autophosphorylation of Hck. Hck (1 μ M) was incubated on ice with 0.1 mM ATP and [γ -³²P]ATP. Aliquots were removed at each time point and resuspended into SDS-PAGE sample buffer. Incorporation of ³²P into the enzyme was determined as described in Methods. **c**, Unphosphorylated Hck or C-terminally phosphorylated Hck was added to a kinase reaction without preincubation in ATP (as in **a**).

essential for HIV-1 replication and pathogenicity¹⁶, binds the SH3 domain of Hck with high affinity and specificity ($K_D \approx 250$ nM)^{17,18}, although the biological significance of this interaction is unknown. Addition of HIV-1 NL4-3-derived Nef to C-terminally phosphorylated Hck, or to Hck that had been preactivated by incubation with

ATP, dramatically increases enzyme activity (Fig. 2a). PA1, a mutant of Nef lacking the polyproline motif necessary for Nef-Hck interaction¹⁷, did not activate Hck (Fig. 2a). The K_{act} for Nef, the concentration at which Nef gives half-maximal activation, is 130 ± 18 nM. In contrast to the results for Hck, Nef does not

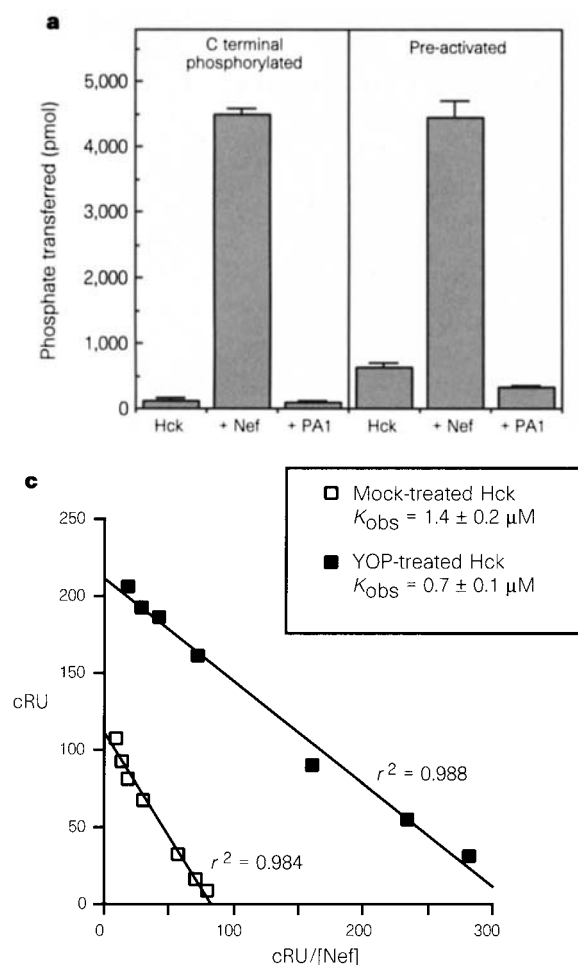


Figure 2 Activation of Hck by HIV-1 Nef. **a**, Nef or PA1 (Nef with Pro $\rightarrow$ Ala mutation in the PXXP motif) ($2 \mu M$) were added to kinase reactions containing peptide substrate together with Hck (C-terminally phosphorylated) or Hck that had been activated by pretreatment with ATP (preactivated). Phosphate transferred to peptide substrate was measured in triplicate after a 10-min reaction, and data are given as mean $\pm$ s.d. **b**, Preactivated Hck was assayed in the presence of various activators and synthetic peptides. Phosphate transferred to peptide substrate was measured in triplicate after a 10-min reaction, and data are given as mean $\pm$ s.d. PYEEI: Glu-Pro-Gln-pTyr-Glu-Glu-Ile-Pro-Ile-Tyr-Leu; SH3 pep: Ser-Pro-Pro-Thr-Pro-Lys-Pro-Arg-Pro-Pro-Arg-Pro. All added reagents were at a final concentration of $2 \mu M$, except for the SH3-binding peptide, which was at 1.0 mM. The affinity of the SH3-binding peptide was not measured, but is expected to be in the 5 – $50 \mu M$ range²⁴; concentrations of this peptide below $100 \mu M$ gave no activation. Addition of higher concentrations of pYEEI showed no additional activation of Hck. **c**, Determination of the equilibrium dissociation constants for interactions between Nef and dephosphorylated Hck (YOP-Hck) or C-terminally phosphorylated Hck (Mock-Hck) by SPR. Nef protein at various concentrations was added to buffer flowing over a biosensor chip coated with either YOP- or mock-treated Hck. The corrected response (cRU) values were determined as described in Methods, and the data were analysed on Scatchard plots. The Scatchard analysis from one such SPR experiment is shown ($n = 3$). The lines that best fit the data are shown together with their coefficients of determination (r^2). Here, the derived dissociation constants are $0.67 \mu M$ for YOP Hck and $1.34 \mu M$ for MOCK Hck. Average values for K_{obs} calculated using data from all three experiments are shown in the inset.

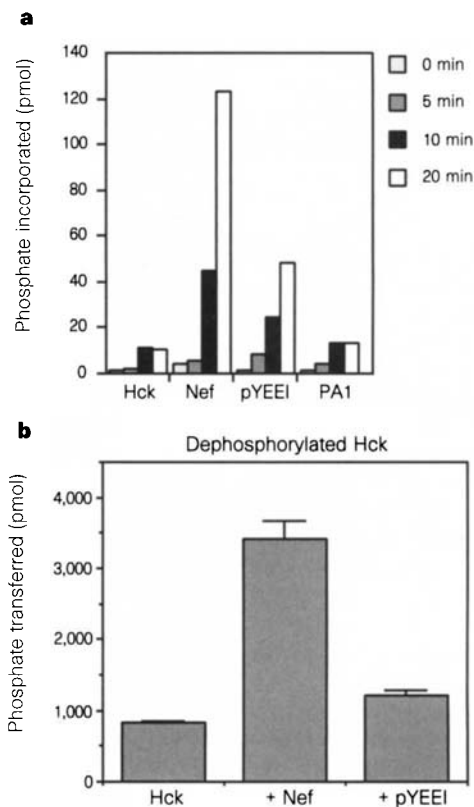


Figure 3 Mechanism of Nef activation. **a**, Hck autophosphorylation (as for Fig. 1c) was measured in the presence of Nef, PA1 or pYEEI peptide (each at $2 \mu M$). Incorporation of ^{32}P into Hck was measured after 0, 5, 10 and 20 min of reaction. **b**, Activation of the unphosphorylated form of Hck was measured in the presence of Nef or pYEEI peptide (each at $2 \mu M$).

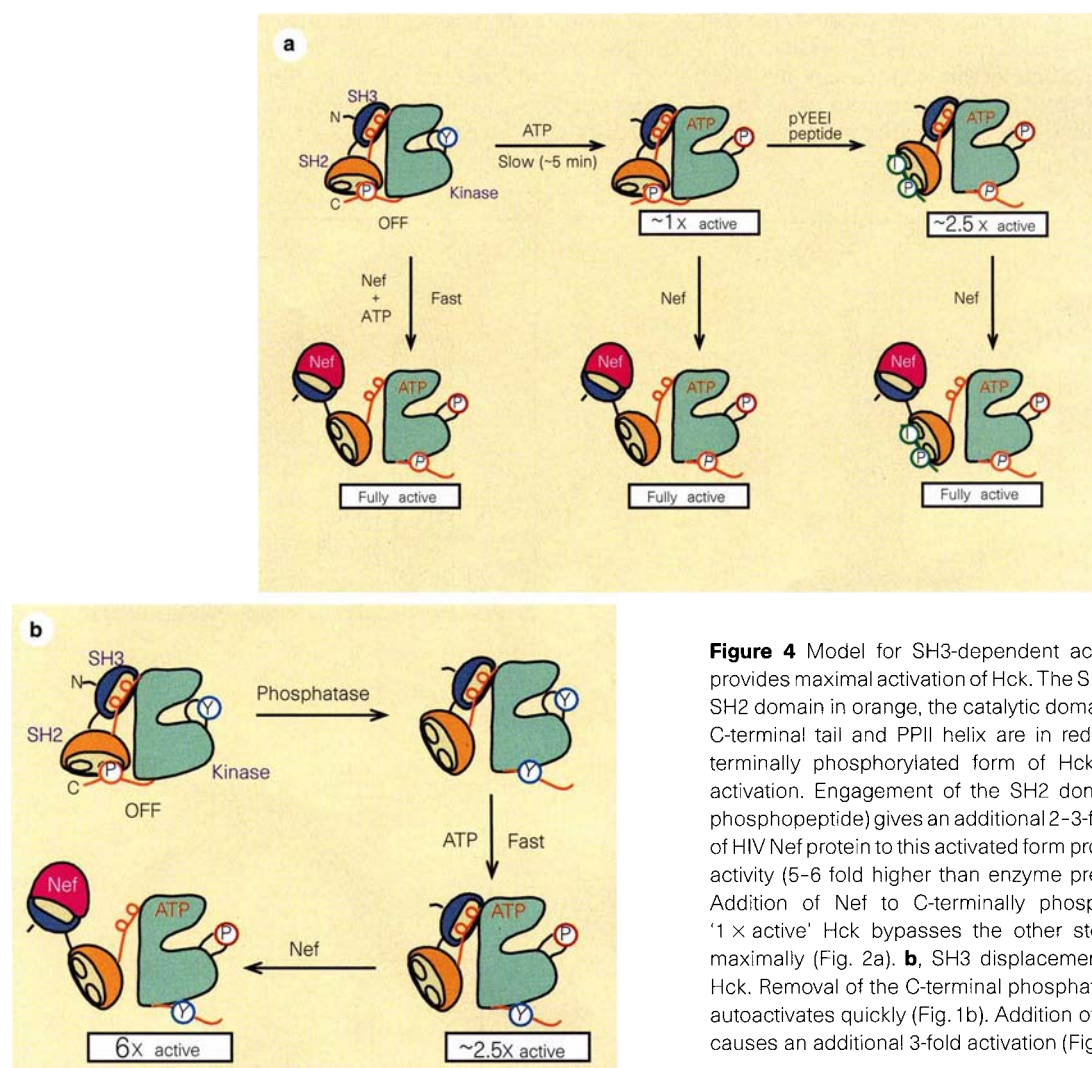


Figure 4 Model for SH3-dependent activation. **a**, SH3 displacement provides maximal activation of Hck. The SH3 domain is shown in blue, the SH2 domain in orange, the catalytic domain in green, Nef in red, and the C-terminal tail and PPII helix are in red-orange. Incubation of the C-terminally phosphorylated form of Hck with ATP produces a slow activation. Engagement of the SH2 domain (for example by a pYEEI phosphopeptide) gives an additional 2–3-fold activation (Fig. 2b). Addition of HIV Nef protein to this activated form provides the maximum enzymatic activity (5–6 fold higher than enzyme pretreated with ATP) (Fig. 2a, b). Addition of Nef to C-terminally phosphorylated Hck ('OFF') or to '1 × active' Hck bypasses the other steps and activates the kinase maximally (Fig. 2a). **b**, SH3 displacement activates dephosphorylated Hck. Removal of the C-terminal phosphate produces a form of Hck that autoactivates quickly (Fig. 1b). Addition of Nef to dephosphorylated Hck causes an additional 3-fold activation (Fig. 3b).

activate the Src family kinase Lck (data not shown); the SH3 domain of Lck does not bind to the PXXP motif of Nef¹⁷. A 12-residue peptide containing the PXXP motif and flanking sequences that is a ligand for Src-family SH3 domains¹⁹ was used for comparison: this peptide also increased the activity of Hck (Fig. 2b). As reported for c-Src¹¹, a phosphopeptide ligand for the Hck SH2 domain increased Hck activity (Fig. 2b), although Nef is a more effective activator under these conditions. We further investigated the influence of SH2-domain occupancy on SH3 availability by comparing the binding of Nef to C-terminally phosphorylated and dephosphorylated Hck using surface plasmon resonance. The C-terminally phosphorylated form of Hck bound Nef with an apparent equilibrium dissociation constant (K_{obs}) of 1.4 μ M, whereas dephosphorylated Hck gave a K_{obs} of 0.7 μ M under the same conditions (Fig. 2c). Thus, release of the C-terminal phosphate from the SH2 domain of Hck may increase access of Nef to the enzyme's SH3 domain.

We tested whether binding of Nef to Hck promotes normal autophosphorylation. Addition of Nef to Hck increased the rate and total amount of kinase autophosphorylation, whereas the PA1 mutant had no effect (Fig. 3a). To test how binding of Nef to autophosphorylated Hck causes increased activation of the enzyme (Fig. 2a), we carried out experiments on dephosphorylated Hck, where the intramolecular SH2–Tyr 527 interaction is no longer possible. Nef activates the dephosphorylated form of Hck ~3-fold, whereas the pYEEI peptide did not stimulate this form of Hck significantly (Fig. 3b). Thus, binding of Nef to the Hck SH3 domain

promotes autophosphorylation, but also activates the tyrosine kinase by a different mechanism. The structure of Hck¹² suggests how this might happen: the SH3 domain docks onto a polyproline type II (PP-II) helix that is formed by the polypeptide chain linking the SH2 domain to the catalytic domain. This linker positions the side chain of Trp 260 so that it interacts with side chains presented by a distorted configuration of helix α C in the amino-terminal lobe of the catalytic domain. This distorted configuration is also stabilized by interactions with the unphosphorylated activation segment of Hck (containing Tyr 416). In addition, the structure of the regulatory apparatus (SH3, SH2, PP-II linker) suggests that flexibility of the catalytic domain would be impeded. The stimulation of Hck by Nef is consistent with a model in which release of the SH3–PPII linker interaction destabilizes the inactive conformation of α C, allowing the activation segment to adopt a conformation competent for phosphotransfer²⁰ (Fig. 4). This mode of activation has the potential to increase Hck kinase activity to a level above that achieved by the conformational switch driven by binding of phosphotyrosine to the SH2 domain (Fig. 3b). A p130^{CAS}-related protein that binds the Src SH3 domain has been demonstrated to raise c-Src activity²¹. Activation of Src family kinases by high-affinity SH3 domain ligands may be a general mechanism for enzyme regulation. □

Methods

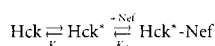
Protein expression and purification. The downregulated form of Hck was expressed and purified as described¹². To prepare dephosphorylated Hck,

purified Hck protein (5 mg) was incubated in 10 mM HEPES, 1 mM DTT, 150 mM KCl, pH 7.5, with 100 units of *Yersinia* protein tyrosine phosphatase (YOP; New England Biolabs) for 5 h at 25 °C. The protein was loaded on a 1-ml Mono-Q column in 20 mM Tris, pH 8.5, 10% glycerol, 1 mM DTT, 1 mM EDTA and eluted with a linear KCl gradient. Fractions containing Hck were pooled, concentrated, and stored at -80 °C in buffer containing 40% glycerol. A mock-treated control sample was prepared the same way except that the YOP enzyme was omitted. The specific activity of this mock-treated sample was comparable to that of untreated Hck. Dephosphorylation was assessed by anti-phosphotyrosine western blotting using antibody PY20 (Transduction Laboratories). HIV-1 NL4-3 Nef protein was expressed and purified as described¹⁷.

Protein kinase assays. Kinase activity assays were performed at 30 °C in buffer containing 20 mM Tris, pH 7.5, 10 mM MgCl₂, 800 μM peptide substrate (RRLLDAHYAARG), [γ -³²P]ATP (100–500 c.p.m. pmol⁻¹) and 500 μM ATP. This peptide is an efficient substrate for Src family kinases (K_m = 240 μM for Src; M.L.-B. and W.T.M., unpublished observations). Reactions were terminated by the addition of cold 10% TCA. The reaction mixtures were centrifuged and aliquots of the supernatants were spotted onto phosphocellulose paper, as described^{22,23}. The phosphocellulose pads were washed and counted in a liquid scintillation counter to measure incorporation of ³²P into peptide. All activity assays were carried out in triplicate. To study activation of C-terminally phosphorylated Hck by ATP, enzyme (10 μM) was incubated on ice for 45 min in various concentrations of ATP. An aliquot of each sample was added to a kinase reaction containing synthetic peptide (final enzyme concentration was 0.44 μM), and the initial rates of phosphotransfer were measured in triplicate. The activation constant for Nef was determined by weighted nonlinear least-squares fit to a hyperbolic velocity versus Nef concentration plot using the iterative program NFIT (Island Products, Galveston, TX). Activators were added at 2 μM unless otherwise noted. The SH2-binding peptide (Glu-Pro-Gln-pTyr-Glu-Ile-Pro-Ile-Tyr-Leu) and SH3-binding peptide (Ser-Pro-Pro-Thr-Pro-Lys-Pro-Arg-Pro-Arg-Pro) were prepared using an Fmoc synthetic strategy, purified by reverse-phase HPLC, and characterized by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry.

Autophosphorylation reactions were carried out on ice in 100 μM ATP. Reactions were terminated with 2 × Laemmli sample buffer and separated on 8% SDS-PAGE. Gels were stained with Coomassie blue and destained for 15 min, and Hck bands were excised and counted in a scintillation counter. Time zero corresponds to an aliquot of the reaction removed after the addition of labelled ATP but before the addition of unlabelled ATP. The site of autophosphorylation was mapped by MALDI-TOF mass spectrometry of tryptic fragments of the enzyme before and after activation.

Surface plasmon resonance (SPR) experiments. These were carried out using a BIAcore Biosensor apparatus (Pharmacia Biosensor). Hck proteins (dephosphorylated with YOP or mock-treated controls) were biotinylated and immobilized on Biosensor SA5 chips as described¹⁸. In all experiments, the equilibrium background response (bulk refractive index) was subtracted from the values obtained from specific binding reactions to yield a specific response value, expressed in corrected response units (cRU). The data were then subjected to Scatchard analysis to determine the dissociation constants. The Nef concentrations used for the Scatchard analysis ranged from 0.1 to 10 μM. All experiments were done in HBS buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.05% surfactant P20) with 1 mM DTT. If the following model for the Hck–Nef interaction is assumed:



where Hck and Hck* represent Hck conformations with an inaccessible or accessible SH3 domain, respectively, K_d is the equilibrium dissociation constant for the Nef–SH3 interaction, and K_c the constant that governs the equilibrium conformation of Hck, then we can say that $[\text{Hck}]_{\text{tot}} - [\text{Hck}^* \cdot \text{Nef}] = (K_{\text{obs}}[\text{Hck}^* \cdot \text{Nef}]/[\text{Nef}])$, where the apparent equilibrium dissociation constant $K_{\text{obs}} = ((1 + K_c)/K_c)K_d$ can be determined by Scatchard analysis. K_{obs} serves as an indicator for the accessibility of the SH3 domain for interaction with other ligands. The linearity of the Scatchard analyses is consistent with the proposed model.

Received 11 December 1996; accepted 22 January 1997.

1. Ziegler, S. F., Marth, J. C., Lewis, D. B. & Perlmutter, R. M. Novel protein-tyrosine kinase gene (*hck*) preferentially expressed in cells of hematopoietic origin. *Mol. Cell. Biol.* **7**, 2276–2285 (1987).
2. Quintrell, N. et al. Identification of a human gene (HCK) that encodes a protein-tyrosine kinase and is expressed in hemopoietic cells. *Mol. Cell. Biol.* **7**, 2267–2275 (1987).
3. Cooper, J. A. & King, C. S. Dephosphorylation or antibody binding to the carboxy terminus stimulates pp60^{src}. *Mol. Cell. Biol.* **6**, 4467–4477 (1986).
4. Courtneidge, S. A. Activation of the pp60^{src} kinase by middle T antigen binding or by dephosphorylation. *EMBO J.* **4**, 1471–1477 (1985).
5. Kmiecik, T. E. & Shalloway, D. Activation and suppression of pp60^{src} transforming ability by mutation of its primary sites of tyrosine phosphorylation. *Cell* **49**, 65–73 (1987).
6. Pivnick-Worms, H., Saunders, K. B., Roberts, T. M., Smith, A. E. & Cheng, S. H. Tyrosine phosphorylation regulates the biochemical and biological properties of pp60^{src}. *Cell* **49**, 75–82 (1987).
7. Cartwright, C. A., Eckhart, W., Simon, S. & Kaplan, P. L. Cell transformation by pp60^{src} mutated in the carboxy-terminal regulatory domain. *Cell* **49**, 83–91 (1987).
8. Cooper, J. A. & Howell, B. The when and how of Src regulation. *Cell* **93**, 1051–1054 (1993).
9. Superti-Furga, G. & Courtneidge, S. A. Structure–function relationships in Src family and related protein tyrosine kinases. *BioEssays* **17**, 321–330 (1995).
10. Roussel, R. R., Brodeur, S. R., Shalloway, D. & Laudano, A. P. Selective binding of activated pp60^{src} by an immobilized synthetic phosphopeptide modeled on the carboxyl terminus of pp60^{src}. *Proc. Natl Acad. Sci. USA* **88**, 10696–10700 (1991).
11. Liu, X. et al. Regulation of c-Src tyrosine kinase activity by the Sec SH2 domain. *Oncogene* **8**, 1119–1126 (1993).
12. Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src-family tyrosine kinase Hck. *Nature* **385**, 602–609 (1997).
13. Cooper, J. A. & MacAuley, A. Potential positive and negative autoregulation of pp60^{src} by intermolecular autophosphorylation. *Proc. Natl Acad. Sci. USA* **85**, 4232–4236 (1988).
14. Barker, S. C. et al. characterization of pp60^{src} tyrosine kinase activities using a continuous assay: autoactivation of the enzyme is an intermolecular autophosphorylation process. *Biochemistry* **34**, 14843–14851 (1995).
15. Boerner, R. J. et al. Correlation of the phosphorylation states of pp60^{src} with tyrosine kinase activity: the intramolecular pY530-SH2 complex retains significant activity if Y419 is phosphorylated. *Biochemistry* **35**, 9519–9525 (1996).
16. Trono, D. HIV accessory proteins: leading roles for the supporting cast. *Cell* **82**, 189–192 (1995).
17. Saksela, K., Cheng, G. & Baltimore, D. Proline-rich (PXXP) sequences in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef⁺ viruses but not for down-regulation of CD4. *EMBO J.* **14**, 484–491 (1995).
18. Lee, C.-H. et al. A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. *EMBO J.* **14**, 5006–5015 (1995).
19. Pleiman, C. M., Hertz, W. M. & Cambier, J. C. Activation of phosphatidylinositol-3' kinase by Src-family kinase SH3 binding to the p85 subunit. *Science* **263**, 1609–1612 (1994).
20. Johnson, L. N., Noble, M. E. M. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149–158 (1996).
21. Alexandropoulos, K. & Baltimore, D. Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel, p130^{cas}-related protein, Sin. *Genes Dev.* **10**, 1341–1355 (1996).
22. Till, J. H., Annan, R. S., Carr, S. A. & Miller, W. T. Use of synthetic peptide libraries and phosphopeptide-selective mass spectrometry to probe protein kinase substrate specificity. *J. Biol. Chem.* **269**, 7423–7428 (1994).
23. Casnellie, J. H. Assay of protein kinases using peptides with basic residues for phosphocellulose binding. *Meth. Enzymol.* **200**, 115–120 (1991).
24. Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3–ligand interactions. *Science* **266**, 1241–1247 (1994).

Acknowledgements. We thank K. Saksela for the original Nef and Hck clones. This work was supported by a grant from the NIH (W.T.M.). F.S. is a fellow of the Human Frontiers Science Program.

Correspondence and requests for materials should be addressed to W.T.M. (e-mail: miller@pofvax.pnb.sunysb.edu).

Interaction between the *C. elegans* cell-death regulators CED-9 and CED-4

Mona S. Spector*, Serge Desnoyers*, Daniel J. Hoepfner† & Michael O. Hengartner*

* Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA

† Graduate Program in Genetics, Department of Molecular Microbiology, SUNY at Stony Brook, Stony Brook, New York 11794, USA

Programmed cell death (apoptosis) is an evolutionarily conserved process used by multicellular organisms to eliminate cells that are not needed or are potentially detrimental to the organism^{1,2}. Members of the Bcl-2 family of mammalian proteins are intimately involved in the regulation of apoptosis, but, their precise mechanism of action remains unresolved^{3–5}. In *Caenorhabditis elegans*, the Bcl-2 homologue CED-9 prevents cell death by antagonizing the death-promoting activities of CED-3, a

member of the Caspase family of death proteases, and of CED-4, a protein with no known mammalian homologue⁶⁻⁹. Here we show that CED-9 interacts physically with CED-4. Mutations that reduce or eliminate CED-9 activity also disrupt its ability to bind CED-4, suggesting that this interaction is important for CED-9 function. Thus, CED-9 might control *C. elegans* cell death by binding to and regulating CED-4 activity. We propose that mammalian Bcl-2 family members might control apoptosis in a similar way through interaction and regulation of CED-4 homologues or analogues.

Genetic analysis of programmed cell death in the nematode *C. elegans* has led to the identification of over a dozen cell-death (*ced*) genes^{10,11}. Two of these genes, *ced-3* and *ced-4*, are required for cells to die: in animals carrying a mutation in either gene, all programmed cell death is blocked¹². The *ced-3* gene encodes a homologue of the mammalian Caspase family of death proteases^{9,13}. *ced-4* is predicted to encode, through alternative splicing, two related but distinct proteins: the major, shorter isoform (CED-4_S) promotes apoptosis⁸, whereas a less abundant, longer isoform (CED-4_L) antagonizes cell death¹⁴. Neither isoform shares any strong sequence similarity with previously characterized proteins^{8,11,15}. A third *C. elegans* gene, *ced-9*, protects *C. elegans* cells from apoptotic death^{6,7,16}. Overexpression of *ced-9* results in the survival of cells that should die; a similar phenotype is also observed in animals carrying the dominant gain-of-function mutation *ced-9(n1950)*. In contrast, *ced-9* loss-of-function mutations result in the widespread death of cells that would normally survive. The CED-9 protein is a member of the Bcl-2 family of cell-death regulators, and Bcl-2 can prevent cell death in *C. elegans* by partially substituting for CED-9 (refs 7,17). The conservation in sequence and function of the CED-3/Caspase and CED-9/Bcl-2 protein families suggests that *C. elegans* and mammals use a similar molecular program to regulate apoptosis^{5,10}.

In *C. elegans*, genetic studies indicate that CED-9 prevents cell death by antagonizing the death-promoting activities of CED-3 and CED-4 (refs 6, 7, 16, 18). To determine whether this suppressive activity might be mediated through a direct interaction, we investigated whether these proteins interact in the yeast two-hybrid system¹⁹. We found that full-length CED-9 binds strongly to CED-4_S, the death-promoting form of CED-4 (Fig. 1 and Table 1). This interaction, which could be detected using either CED-9 or CED-4_S as the bait, is specific, as neither LexA–CED-9 nor LexA–CED-4_S interact with other proteins used as negative controls (GAL4_{AD}–STE11 and GAL4_{AD}–Lamin; Fig. 1 and Table 1, and data not shown).

To determine whether this CED-9/CED-4 interaction in our yeast two-hybrid system was direct, we tested CED-9 and CED-4 for interaction *in vitro*. We found that *in vitro*-translated CED-4_S protein readily bound to a glutathione-S-transferase–CED-9 (GST–CED-9) fusion protein (Fig. 2a, b); there was little binding of CED-4_S and GST–CED-9 to GST–p16 and CDK4 (two irrelevant proteins used as negative controls), respectively, confirming that the CED-9/CED-4_S interaction is specific (Fig. 2a, b).

Given the strong genetic link between *ced-9* and *ced-4*, it is likely that the physical interaction between CED-9 and CED-4 is biologically important. We therefore looked for a correlation between the ability of CED-9 to interact with CED-4 and its capacity to prevent cell death. We introduced point mutations into the *ced-9* open reading frame to recreate three CED-9-mutant proteins whose effect on cell death we had previously characterized: CED-9(G169E), the product of the dominant gain-of-function allele *ced-9(n1950)*; CED-9(Y149N), which corresponds to the loss-of-function mutation *ced-9(n1653ts)*; and Q160stop, which results in the truncated protein CED-9(1–159) and corresponds to the loss-of-function mutation *ced-9(n2077)* (refs 6, 7, 16).

The *n1950* gain-of-function mutation results in a glycine-to-glutamate (G169E) substitution in the conserved Bcl-2 homology

Table 1 Mutational analysis of the CED-9/CED-4 interaction in the yeast two-hybrid system

Bait (LexA fusion)	Interactor (GAL4 _{AD} fusion)	Colour intensity (filters)	Relative β -galactosidase activity
STE11	STE11	+++	4,100 $\pm$ 400
CED-4 _S	CED-9	+++	18.7 $\pm$ 1.5
CED-4 _S	STE11	+/-	3.6 $\pm$ 1.4
CED-9	CED-4 _S	+++	541 $\pm$ 29
CED-9	STE11	-	1.0 $\pm$ 0.3
CED-4 _S	CED-9(1-231)	+++	1,080 $\pm$ 320
STE11	CED-9(1-231)	-	0.6 $\pm$ 0.1
CED-4 _S	CED-9(G169E) (<i>gf</i>)	+++	42.4 $\pm$ 4.2
CED-9(G169E) (<i>gf</i>)	CED-4 _S	+++	455 $\pm$ 93
CED-9(G169E) (<i>gf</i>)	STE11	-	0.8 $\pm$ 0.1
STE11	CED-9(G169E) (<i>gf</i>)	-	0.9 $\pm$ 0.1
CED-4 _S	CED-9(Y149N) (<i>lf</i>)	-	0.9 $\pm$ 0.3
CED-4 _S	CED-9(1-159) (<i>lf</i>)	-	0.6 $\pm$ 0.2
STE11	CED-9(Y149N) (<i>lf</i>)	-	0.5 $\pm$ 0.1
STE11	CED-9(1-159) (<i>lf</i>)	-	0.3 $\pm$ 0.1

Interaction strength was determined both on filters and in liquid^{26,27}. Intensity of the blue colour on filters was assessed on a qualitative scale, from '-' (white) to '+' + '+' (as strong as the LexA–STE11/GAL4_{AD}–STE11 positive control). β -Galactosidase activities shown correspond to average $\pm$ s.e.m. of three independent transformants assayed in duplicate, and are shown as relative activities compared to the negative control (LexA–CED-9/GAL4_{AD}–STE11). All fusion proteins were stably expressed, as determined by western blotting (data not shown). *gf*, Gain-of-function; *lf*, loss-of-function.

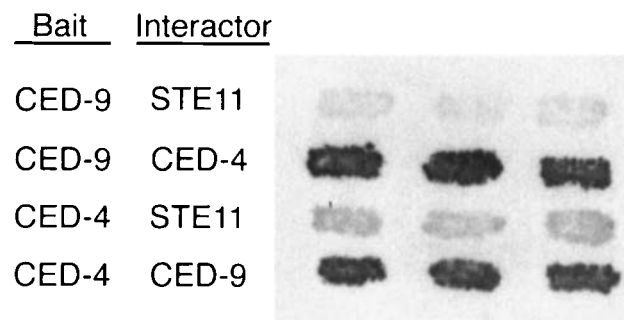


Figure 1 CED-9 and CED-4 interact in the yeast two-hybrid system. LexA and GAL4_{AD} constructs were co-transformed into yeast strain L40 and colonies selected on SD medium lacking leucine and tryptophan. The interaction between the fusion proteins was detected by the induction of β -galactosidase (dark colour, assayed according to ref. 27). Each patch represents an independent yeast transformant.

(BH) domain BH1, which is present in and mediates interactions between most Bcl-2 family members²⁰. Consistent with the observation that *n1950* does not eliminate *ced-9* function, we found that CED-9(G169E) interacts with CED-4_S (Table 1 and Fig. 2c). Because the CED-9(G169E)/CED-4_S interaction was not significantly stronger than the one we observed between CED-9(+) and CED-4_S, *n1950* is unlikely to cause a gain of function by simply increasing CED-9/CED-4 interaction. Rather, we suspect—on the basis of the observation that human Bcl-2 can interact with many distinct proteins—that *n1950* affects the ability of CED-9 to interact with another, as-yet unidentified partner, possibly the nematode equivalent of Bax. Alternatively, *n1950* might alter the *in vivo* regulation of the CED-9/CED-4 interaction in ways that cannot be detected in our assay systems.

In contrast, the Y149N and Q160stop loss-of-function mutations both abrogated interaction between CED-9 and CED-4 in the yeast

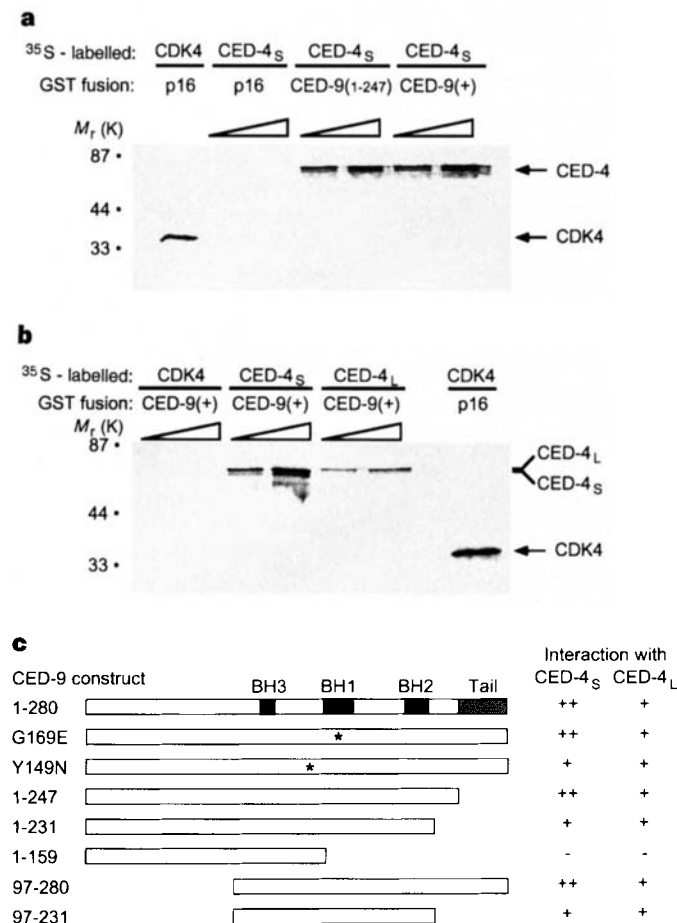


Figure 2 Identification of the CED-9 domain required for interaction with CED-4 *in vitro*. **a**, **b**, CED-4_S and CED-4_L bind to GST-CED-9. Radiolabelled full-length *ced-4_S* and *ced-4_L* ORFs were incubated with increasing amounts (twofold in **a** and tenfold in **b**) or immobilized GST fusion proteins produced in *E. coli*. Bound proteins were analysed by SDS-PAGE, followed by autoradiography. **c**, Interaction between CED-9 truncations and CED-4 isoforms. *ced-9* fragments fused in-frame to GST were tested for interaction with CED-4 isoforms as described in Methods. Relative strength of interaction was determined by comparing the intensities of radiolabelled bands on autoradiograms (+ + +, ++, +, and - correspond to the recovery of about 10, 2-5, 0.2-1.0 and <0.05% of the input radiolabelled protein, respectively; the CDK4/GST-p16 positive control scored a + + + in our assay). The positions of the BH domains²⁹ (black boxes) and C-terminal hydrophobic tail (grey box) are indicated. Asterisk, position of the point mutation in G169E and Y149N constructs.

two-hybrid system (Table 1). Q160stop also eliminated the CED-9/CED-4 interaction *in vitro*, whereas Y149N only reduced it (Fig. 2c). The weak effect of Y149N on CED-9/CED-4 interaction *in vitro* is consistent with the temperature-sensitive nature of this mutation⁶, and the different temperatures at which our assays were performed (30 °C for yeast; 4 °C *in vitro*). *In vivo*, Y149N might prevent CED-9/CED-4 interaction only at high but not at low temperatures (we have excluded trivial explanations, such as reduced protein stability for the n1653 temperature sensitivity, as the Y149N protein is as stable as the wild-type protein under both permissive and restrictive temperatures; M.S.S., unpublished results). The correlation between loss of *ced-9* function and loss of CED-4 binding is consistent with the hypothesis that direct interaction with CED-4 is important for CED-9 to prevent cell death.

To identify the domain(s) of CED-9 required for interaction with CED-4, we generated several CED-9-deletion constructs and tested

their ability to interact with full-length CED-4_S *in vitro* (Fig. 2c). Several Bcl-2 family members contain a C-terminal hydrophobic tail, which drives these proteins to the outer surface of the mitochondrion, the rough endoplasmic reticulum and nuclear membranes²¹⁻²³. Truncations that remove this tail only slightly reduce the ability of Bcl-2 to prevent apoptosis^{24,25}. Similarly, we found that CED-9(1-247), which lacks the hydrophobic tail, still prevents programmed cell death when overexpressed in *C. elegans* (data not shown). Thus, the C-terminal tail is not essential for Bcl-2 or CED-9 function. Consistent with these observations, CED-9(1-231), which lacks the C-terminal 49 amino acids, still interacted with CED-4_S (Table 1 and Fig. 2c). However, elimination of a further 72 amino acids, which include the BH1 and BH2 domains (CED-9(1-159) versus CED-9(1-231)), completely prevented CED-9 from interacting with CED-4_S both *in vitro* and in yeast (Table 1 and Fig. 2c), indicating that the BH1 and BH2 domains, which are important for the function of several Bcl-2 family members, might be involved in CED-9/CED-4 interaction. In support of this, residues 97-231 (containing the BH1, BH2 and BH3 domains) were sufficient for interaction with CED-4_S (Fig. 2c).

An alternative splice variant of *ced-4*, *ced-4_L*, has been described¹⁴. This variant encodes a putative protein containing 24 additional amino acids. In contrast to the principal isoform (*ced-4_S*), overexpression of *ced-4_L* prevents death. However, both forms of *ced-4* are antagonized by *ced-9* (ref. 14), possibly explaining why *ced-9* also has a minor death-promoting activity¹⁶. Because *ced-9* antagonizes both *ced-4_L* and *ced-4_S*, we suspected that CED-9 could also interact with CED-4_L. Indeed, we found no qualitative differences between CED-4_S and CED-4_L in their ability to bind to CED-9, although the CED-9/CED-4_L interaction was invariably weaker (Fig. 2b, c). Thus, the difference between CED-4_S and CED-4_L may be in their ability to interact with or activate downstream targets, rather than in their binding to CED-9.

We have demonstrated that the *C. elegans* death suppressor CED-9 interacts directly with one of its downstream targets, CED-4. This interaction is disrupted in two *ced-9(lf)* mutants and might be mediated through the conserved BH1, BH2 and BH3 domains. Although our data indicate that binding to CED-4 is important for CED-9 function, we do not know whether it is sufficient. In a simple model consistent with our observations and with the genetic data, CED-9-bound CED-4_S would be unable to promote apoptosis and death-promoting stimuli would induce dissociation of the two proteins. The conservation of the death machinery between nematodes and mammals indicates that, as for *ced-3* and *ced-9*, mammals should contain one or several proteins that play a role similar to that of *ced-4* in *C. elegans*. If the CED-4-binding domain in CED-9 is conserved through evolution, then this information could be used to identify mammalian CED-4 homologues (or analogues) through their ability to interact with the equivalent domain in Bcl-2 family members.

Methods

Site-directed mutagenesis and deletion constructs. The n1653 mutation was introduced into a *ced-9* open reading frame (ORF) construct by site-directed mutagenesis (Quik Change, Stratagene), as recommended by the manufacturer. The n1950 ORF was generated as described¹⁶. All deletion constructs were generated by polymerase chain reaction (PCR) using appropriate primers. To obtain a full-length *ced-4_L* clone, PCR was used with reverse transcription to amplify (from total, mixed-stage RNA) a cDNA fragment covering the alternatively spliced exon¹⁴; this *ced-4_L* fragment was then used to replace the corresponding region in *ced-4_S* using unique restriction sites present in the *ced-4* ORF. The DNA sequence of all constructs used were confirmed by fluorescent sequencing on an ABI 377 Sequencer.

Yeast two-hybrid assay. *ced-9* and *ced-4* cDNAs and fragments were cloned in-frame into the yeast two-hybrid vectors pBTM116 (to create LexA fusions) and pGAD-GH (to create GAL4_{AD} fusions)^{26,27}. LexA and GAL4_{AD} constructs were co-transformed into yeast strain L40 and colonies selected on SD medium

lacking leucine and tryptophan. LexA-STE11, GAL4_{AD}-STE11, and GAL4_{AD}-lamin constructs used in control experiments were generously provided by L. Van Aelst. β -Galactosidase was assayed as described²⁸.

In vitro interaction. The *ced-9* open reading frame and various deletion fragments were subcloned in-frame into the GST fusion construct pGEX-1XT (Pharmacia). GST fusion constructs were transformed into *Escherichia coli* strain BL21(DE3) and induced by addition of IPTG. Fusion proteins were purified using glutathione-Sepharose beads (Pharmacia), as suggested by the manufacturer. Full-length *ced-4*_S and *ced-4*_L ORFs in pCITE (Novagen) were used as templates in an *in vitro* transcription/translation reaction (TnT, Promega) in the presence of ³⁵S-methionine. Pre-cleared *in vitro* transcription/translation lysates were added to 200 μ l incubation buffer (150 mM NaCl, 50 mM Tris, pH 8.0, 2 mM EDTA, 5 mM DTT, 0.5% Nonidet P-40) containing the GST fusion protein (between 0.2 and 2.0 μ g, depending on the experiment) immobilized on glutathione-Sepharose, incubated for 1 h at 4 °C, and washed four times with incubation buffer. Bound proteins were analysed by SDS-PAGE followed by autoradiography.

Received 20 January; accepted 28 January 1997.

1. Ellis, R. E., Yuan, J. & Horvitz, H. R. Mechanisms and functions of cell death. *Annu. Rev. Cell Biol.* **7**, 663–698 (1991).
2. Thompson, C. B. Apoptosis in the pathogenesis and treatment of disease. *Science* **267**, 1456–1462 (1995).
3. Hacker, G. & Vaux, D. L. A sticky business. *Curr. Biol.* **5**, 622–624 (1995).
4. Korsmeyer, S. J. Regulators of cell death. *Trends Genet.* **11**, 101–105 (1995).
5. Chinnaiyan, A. M. & Dixit, V. M. The cell-death machine. *Curr. Biol.* **6**, 555–562 (1996).
6. Hengartner, M. O., Ellis, R. E. & Horvitz, H. R. *C. elegans* gene *ced-9* protects cells from programmed cell death. *Nature* **356**, 494–499 (1992).
7. Hengartner, M. O. & Horvitz, H. R. *C. elegans* cell death gene *ced-9* encodes a functional homolog of mammalian proto-oncogene *bcl-2*. *Cell* **76**, 665–676 (1994).
8. Yuan, J. & Horvitz, H. R. The *Caenorhabditis elegans* cell death gene *ced-4* encodes a novel protein and is expressed during the period of extensive programmed cell death. *Development* **116**, 309–320 (1992).
9. Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M. & Horvitz, H. R. The *C. elegans* cell death gene *ced-3* encodes a protein similar to mammalian interleukin-1 β converting enzyme. *Cell* **75**, 641–652 (1993).
10. Hengartner, M. O. & Horvitz, H. R. Programmed cell death in *C. elegans*. *Curr. Opin. Gen. Dev.* **4**, 581–586 (1994).
11. Horvitz, H. R., Shaham, S. & Hengartner, M. O. The genetics of programmed cell death in the nematode *Caenorhabditis elegans*. *Cold Spring Harbor Symp. Quant. Biol.* **LIX**, 377–385 (1994).
12. Ellis, H. M. & Horvitz, H. R. Genetic control of programmed cell death in the nematode *C. elegans*. *Cell* **44**, 817–829 (1986).

13. Alnemri, E. S. *et al.* Human ICE/CED-3 protease nomenclature. *Cell* **87**, 171 (1996).
14. Shaham, S. & Horvitz, H. R. An alternatively spliced *C. elegans ced-4* RNA encodes a novel cell death inhibitor. *Cell* **86**, 201–208 (1996).
15. Nagata, S. Telling cells their time is up. *Curr. Biol.* **6**, 1241–1243 (1996).
16. Hengartner, M. O. & Horvitz, H. R. Activation of *C. elegans* cell death protein CED-9 by an amino-acid substitution in a domain conserved in Bcl-2. *Nature* **369**, 318–320 (1994).
17. Vaux, D. L., Weissman, I. L. & Kim, S. K. Prevention of programmed cell death in *Caenorhabditis elegans* by human *bcl-2*. *Science* **258**, 1955–1957 (1992).
18. Shaham, S. & Horvitz, H. R. Developing *Caenorhabditis elegans* neurons may contain both cell-death protective and killer activities. *Genes Dev.* **10**, 578–591 (1996).
19. Chien, C. T., Bartel, P. L., Sternglanz, R. & Fields, S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl Acad. Sci. USA* **88**, 9578–9582 (1991).
20. Farrow, S. & Brown, R. New members of the Bcl-2 family and their protein partners. *Curr. Opin. Genet. Dev.* **6**, 45–49 (1996).
21. Chen-Levy, Z., Nourse, J. & Cleary, M. L. The *bcl-2* candidate proto-oncogene product is a 24-kilodalton integral-membrane protein highly expressed in lymphoid cell lines and lymphomas carrying the t(14;18) translocation. *Mol. Cell. Biol.* **9**, 701–710 (1989).
22. Hockenbery, D., Nuñez, G., Millman, C., Schreiber, R. D. & Korsmeyer, S. J. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature* **348**, 334–336 (1990).
23. Krajewski, S. *et al.* Immunohistochemical determination of *in vivo* distribution of Bax, a dominant inhibitor of Bcl-2. *Am. J. Pathol.* **145**, 1323–1336 (1994).
24. Hockenbery, D. M., Oltvai, Z. N., Yin, X.-M., Millman, C. L. & Korsmeyer, S. J. Bcl-2 functions in an antioxidant pathway to prevent apoptosis. *Cell* **75**, 241–251 (1993).
25. Hunter, J. J., Bond, B. L. & Parslow, T. G. Functional dissection of the human Bcl-2 protein: sequence requirements for inhibition of apoptosis. *Mol. Cell. Biol.* **16**, 877–883 (1996).
26. Van Aelst, L., Joneson, T. & Bar-Sagi, D. Identification of a novel Rac 1-interacting protein involved in membrane ruffling. *EMBO J.* **15**, 3778–3786 (1996).
27. Van Aelst, L. Two hybrid analysis of Ras-Raf interactions. In *Methods in Molecular Biology* (Humana, Totowa, NJ, USA, in the press).
28. Miller, J. H. Assay of β -galactosidase. In *Experiments in Molecular Genetics* 352–355 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1972).
29. Muchmore, S. W. *et al.* X-ray and NMR structure of human Bcl-x₁, an inhibitor of programmed cell death. *Nature* **381**, 335–341 (1996).

Acknowledgements. We thank G. Jefferis and T. Casci for reagents, A. Chinnaiyan and V. Dixit for communicating results before publication, G. Hannon and M. Serrano for Cdk4 and GST-p16 constructs and discussions, L. Van Aelst for LexA-lamin, LexA-STE11 and pGAD-STE11 constructs, J. Keller and S. Teplin for technical assistance, and members of our laboratory for critical comments and suggestions. This work was supported by grants from the US Public Health Service and the Donaldson Charitable Trust. M.S.S. was supported by an NIH training grant; S.D. was supported by the GlaxoWellcome Foundation, FRSQ (Québec), and the Human Frontier Science Program; M.O.H. is a Rita Allen Foundation Scholar.

Correspondence and requests for materials should be addressed to M.O.H. (e-mail: hengartn@cshl.org).

KNOW YOUR COPY RIGHTS RESPECT OURS

The publication you are reading is protected by copyright law.
Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Diverse opportunities for immunologists

Close links between courses and senior scientists are one key to career development in immunological fields. And better organization of careers advice is on the way.

Brendan Horton and Orla Smith

The 20 January issue of *Time* magazine featured a rundown of hot fields for job seekers, and listed the emerging 'boom towns' in the United States where many companies are relocating and new businesses are finding their feet. Salt Lake City in Utah, Raleigh in North Carolina and Seattle in Washington state are among a handful of cities experiencing considerable growth in their high-technology job markets.

Peter Linsley moved to Seattle following an academic postdoctoral fellowship at the University of Toronto to join a small biotechnology company that was later taken over by pharmaceutical giant Bristol-Myers Squibb. As a senior research fellow, Linsley is developing monoclonal antibody-fusion proteins that block T-cell activation. He has enjoyed the challenge of working in industry, which he says "has provided the opportunity to publish extensively and to grow as a scientist". Linsley thinks there are plenty of opportunities in industry for immunologists — when it comes to selecting candidates, he says versatility and the ability to think critically are more important than specific areas of expertise.

On the move

Moving between laboratories in different countries can have its benefits. Hidde Ploegh, an immunologist at the Massachusetts Institute of Technology (MIT) who studies antigen processing, says: "Every time I have moved it has proved to be an intellectual rejuvenation". Originally from the Netherlands, Ploegh completed his PhD in Jack Strominger's laboratory at Harvard and moved to the Institute for Genetics in Cologne, Germany, to run his own laboratory. After four years, he was recruited by the Netherlands Cancer Institute to head its department of cellular biochemistry. After eight years he returned to the United States as a tenured full professor at MIT where he now studies host-pathogen inter-relationships.

A useful source of funding for scientists on the move is the Human Frontier Science Program (HFSP), which supports international collaborations, particularly among younger scientists, by providing funds for basic research. Areas of research backed by the programme include the elucidation of brain function and biological functions through molecular approaches, molecular immunology, receptor-ligand interaction, and signal transduction. HFSP grants are available for up to three years for projects involving teams from different countries.

Fellowships are of two types: short-term grants of up to three months' funding; and longer-term grants covering projects of one or two years.

Expertise in science goes well beyond the bench and the classroom these days. M. Michele Hogan, executive director of the American Association of Immunologists (AAI), has succeeded on a course that did not stick to the old standard route of PhD, postdoc, assistant professorship and tenure. After she received a masters degree in biology and anthropology she worked as a physical anthropologist on archaeological excavations in Herzelya, Israel. She went on to receive a PhD in pathobiology from the University of Minnesota, after which she moved to the US National Institutes of Health (NIH) for postdoctoral training and a stint as a research associate.

When it came time to start looking for a 'real job', Hogan considered her options in academic research, government and industry. She visited NIH's National Institute of Allergy and Infectious Diseases (NIAID) in search of openings. "What started as a fact-finding mission turned into a job offer", says Hogan. Subsequently, she became the programme officer in the genetics and transplantation branch.

She held several other positions in the same division before being promoted to chief of the basic immunology branch in the Division of Allergy, Immunology and Transplantation. She became an advocate and a

presenter of applications that had just 'missed the payroll' to NIAID's national advisory council for a second consideration. Other responsibilities included managing the progress of contracts and writing research grant proposals.

In this capacity, Hogan interacted with many (if not most) of the country's most prominent immunologists, as well as with individual trainees, postdocs, first-time grant holders and institutions. She has served on several of NIH's administration and policy committees and programmes, and has represented NIAID in other federal programmes, speaking on NIAID activities of public interest. According to Hogan, this position required "understanding, commitment and training in immunology, as well as a personality that 'enjoyed working with scientists and being a liaison for them'".

Since last September, Hogan has been executive director of AAI, being involved in the policy issues of science, as well as their implementation. As a federal employee she had been prevented from any direct involvement in this kind of work, which therefore made the private sector appealing to her.

Advanced courses

MIT's Ploegh is an enthusiastic educator, teaching on the AAI advanced course in immunology at the University of California, Berkeley, and the NATO biennial immunol-

Watch the Internet for careers update

The American Association of Immunologists (AAI) has about 5,300 members and, according to M. Michele Hogan, executive director, is growing at about 0.5–1 per cent a year. But the number of society members tells little about the job market.

In the past, information on career trends has not been made available in a timely fashion. In 1995, the US National Academy of Science's Committee on Science, Engineering and Public Policy (COSEPUP) published a report entitled *Reshaping the Graduate Education of Scientists and Engineers*. In it, the committee expressed concern over the lack of organization and timeliness of career information, as well as the lack of guidance given to students and careers advisers. It made the National Science Foundation (NSF) responsible for improving matters.

Consistent with the recommendations of COSEPUP, the Commission on Professionals in Science and Technology

was created as a participating organization of the American Association for the Advancement of Science and received a \$415,000 grant from the Alfred P. Sloan Foundation. The project was set up "to formulate and carry out a plan for the continuous monitoring of the supply and demand situation for recent graduates in science and engineering".

Initially the commission began working with seven societies, including the American Chemical Society and the American Institute of Physics. Last October, NSF committed another \$430,000, which has allowed six more societies to become involved, including the American Society for Microbiology and the American Physiological Society. It is hoped that more money will be awarded later this year so that more societies can be added.

Career data will be available from autumn 1997 and will be published on the Internet.

B.H.

Table 1 **Sites of Internet interest for immunologists**

Site name	URLs: http://
Organizations and Societies	
The American Association of Immunologists	glamdring.ucsd.edu/others/aai/
The American Academy of Allergy, Asthma and Immunology	www.aaaai.org
American Federation for Clinical Research	www.afcr.org/index.htm
Commission on Professionals in Science and Technology	www.aaas.org/cpst/
National Institute of Allergy and Infectious Diseases Home Page	www.niaid.nih.gov/
National Jewish Center for Immun. and Resp. Med.	www.njc.org/
The CWRU Immunology Advanced Technology Lab	mediswww.meds.cwrui.edu/dept/ATL/
Microbiology, Immunology and Virology Websites	bioinformatics.weizmann.ac.il/ISICR/
The International Society for Interferon and Cytokine Research	http://bioinformatics.weizmann.ac.il/ISICR/
Molecular Biology Databases - Subject-Specific Databases	bioinformatics.weizmann.ac.il/mb/db/subject_specific_databases.html#immunology
Arthritis Foundation - Research Resources	www.arthritis.org/research/
Federation of American Societies for Experimental Biology	www.faseb.org/careers/
The National Academy of Science	www.nas.edu/
Corporations	
Bristol-Myers Squibb	www.bms.com
CYTimune Sciences	www.cytimmune.com/
Endogen	www.endogen.com/
PharMingen Home Page	www.pharmingen.com/
Santa Cruz Biotechnology	www.scbt.com/
Immunotech	www.immunotech.com/
Education	
University of Adelaide - Microbiology and Immunology	www.microbiology.adelaide.edu.au/
Sackler School Program in Immunology	www.healthsci.tufts.edu/Immunology/imm.html
Immunology Training Program	med-med1.bu.edu/immun.html
Immunology Dept., University of Toronto	www.immune.med.utoronto.ca/
AAI listing of graduate programmes	glamdring.ucsd.edu/others/aai/96prelimprogram.html#glamdring.ucsd.edu/others/aai/gradprog.html
Center for Immunology, University of Minnesota	http://www.borg.labmed.umn.edu/imm.html
Employment	
AAI - faculty openings	glamdring.ucsd.edu/others/aai/postdocs.html glamdring.ucsd.edu/others/aai/facultyopen.html
National Jewish - Human Resources Jobline	
Telephone Service	www.njc.org/HR/HR_jobs.html
Conferences and meetings	
Fifth International Cytokine Conference	bioinformatics.weizmann.ac.il/cytokine/conf_nevada.html
NMHCC and the Bio/Technology Conferences	bioinformatics.weizmann.ac.il/cytokine/conf_NMHCC.html
The AAAAI/AAI/CIS joint meeting preliminary programme	http://glamdring.ucsd.edu/others/aai/96prelimprogram.html
The P/L/S Group conferences by subject	www.pslgroup.com/dg/immuno.htm
The P/L/S Group conferences - immunology and allergy	www.pslgroup.com/dg/immuno.htm
Interleukin-10 Workshop	bioinformatics.weizmann.ac.il/cytokine/conf_glasgow.html
1997 Annual Meeting of the ISICR	http://bioinformatics.weizmann.ac.il/cytokine/conf_sandiego.html
Grants and funding	
National Jewish - Funding Resources on the Web	www.njc.org/LIBRhtml/Grants_info.html
The Illinois Researcher Information Service (IRIS) and OPS	carousel.lis.uiuc.edu/~iris/about_iris.html
The Foundation Center	FDNCENTER.ORG/
US Federal Government Agencies	carousel.lis.uiuc.edu/~iris/federal.html
Funding Opportunities in Health and Safety/Medical Sciences/Biomedical	cos.gdb.org/repos/fund/disc/health.html
FEDIX	web.fie.com/fedix/index.html
Human Frontier Science Program Organization (HFSP)	www.hfsp.org/

ogy course held by the Federation of European Biochemical Societies in Greece. Both courses are designed to teach graduate and postdoctoral students the latest advances in immunobiology. The FEBS NATO course is organized by Diane Mathis, an immunologist who travelled in the opposite direction to Ploegh, completing her PhD at Stanford and moving to the Institut de Genetique et de Biologie Moleculaire et Cellulaire in Strasbourg, France. To teach the eight-day course, Mathis, together with her distinguished colleague Klaus Rajewsky, selects renowned immunologists from the United States and Europe. The course alternates with one organized by the European network of immunology institutes, which this year will be in Portugal.

The AAI supports education through conferences such as the AAAAI/AAI/CIS joint meeting, which is next held later this month (21–26 February) in San Francisco (details are on AAI's home page, see Table 1). The society also supplies a wealth of information on its Web site, with a list of worldwide education forums and conferences in immunology. AAI runs its own advanced course in immunology on 14–19 July at the University of California, Berkeley, similar to the FEBS NATO course. □

Brendan Horton is in Nature's Washington office and Orla Smith is news and views editor of Nature Medicine.

The dos and don'ts of grant writing

Diane Gershon

The US National Institutes of Health (NIH) is making an effort to devote more funds to the support of investigator-initiated research projects. As a result, institutes such as the National Institute of Allergy and Infectious Diseases (NIAID) and the National Cancer Institute (NCI) have over the past couple of years been changing the way they do business.

'Request for application' (RFA grants) tend to be very specific, often pinpointing the strategy or approach to be taken to tackle a problem, which can tie a researcher's hands. They are also usually one-time-only arrangements with stated set-aside funds. And, unlike other grant applications, RFAs are reviewed by the same institute that solicits the research proposals, rather than by an outside study section.

In the case of NIAID, only 15 per cent of funds made available for so-called competing awards are set aside for RFAs and 'request for proposals' (which are contracts). The remaining 85 per cent is devoted to 'unsolicited' research. Nevertheless this subtle shift in emphasis away from targeted 'contract-type' research has been warmly received by investigators.

Grant applications submitted in response to programme announcements (details of

Table 2 **US Sources of funding other than NIH**

Government

National Science Foundation www.nsf.gov

Arlington, Virginia, 202-357-7905. Research grants, fellowships, career/development & salary support

Private

American Cancer Society, Atlanta, Georgia tel. 404-320-3333. Research grants, fellowships, career development awards

American Foundation for AIDS Research, Los Angeles, California tel. 213-857-5900. Research grants & career development awards

American Society of Clinical Oncology, Chicago, Illinois tel. 312-644-0828

Arthritis Foundation, Atlanta, Georgia tel. 404-872-7100. Research grants, new investigator grants, postdoctoral fellowships & career development awards

Burroughs Wellcome Fund, Morrisville, North Carolina tel. 919-991-5100. Career development and developing investigator awards

Cancer Research Institute, New York, New York tel. 212-688-7515. Postdoctoral fellowships & career development awards

Damon Runyon-Walter Winchell Foundation, New York, New York tel. 212-532-3888. Postdoctoral research fellowships

Howard Hughes Medical Institute, Chevy Chase, Maryland tel. 301-215-8500/<http://www.hhmi.org/> Predoctoral fellowships, research training fellowships, research scholars at NIH, postdoctoral fellowships

Immune Deficiency Foundation, Columbia, Maryland tel. 410-461-3127. Grants & postdoctoral fellowships

Infectious Diseases Society of America, Washington, DC tel. 202-857-1139.

Young investigator awards

Irvington Institute, New York, New York tel. 212-758-8968. Postdoctoral fellowships

Leukemia Research Foundation, New York, New York

tel. 212-573-8484. Research grants & postdoctoral fellowships

G. Harold & Leila Y. Mathers Charitable Foundation, Mount Kisco, New York tel. 914-242-0465. Grants

National Leukemia Association, Garden City, New York tel. 516-222-1944. Research grants

Pew Scholars Program, Philadelphia, Pennsylvania tel. 215-575-9050. Career development support

Pfizer Pharmaceuticals, New York, New York tel. 212-573-2880. Pfizer scholars program for new faculty

Other useful contacts

American Association of Immunologists
<http://glandring.ucsd.edu/others/aai>

- Amgen
- Bayer Corporation
- Behringwerke
- Boehringer Ingelheim Pharmaceuticals, Inc.
- Biogen
- Bristol-Myers Squibb Pharmaceutical Research Institute
- DNAX
- Division of Research Grants, NIH
- Endogen
- Entre Med, Inc.
- Genetics Institute, Inc.
- Glaxo
- ImmunLogic
- Pharmaceutical Corporation Immunex Corporation
- Jackson Immunoresearch Labs, Inc.
- Merck & Co, Inc.
- MURO Pharmaceutical, Inc.
- National Cancer Institute, NIH
- National Institute of Allergy and Infectious Diseases, NIH
- Pfizer, Inc.
- Pharmacia Biotech AB
- PharMingen
- Searle
- Systemix
- Vertex

which are published in the *NIH Guide* and can be found on NIH's Internet home page, <http://www.nih.gov> are peer-reviewed by study sections, so avoiding any appearance of favouritism by an institute.

Although it is early days, the effects of this policy shift can already be seen, for example in NCI where funding for RFAs dropped sharply by 33 per cent, from \$30 million in financial year 1995 to less than \$20 million in 1996. The funds were redirected to help boost the pay-line for the institute's RO1 grants.

Most institutes have a way of dealing with research grants on the margin of the pay-line. Some set aside a financial reserve to pay certain grants as 'exceptions'. Stephen M. Hazen, chief of NCI's extramural financial data branch, says: "If we [NCI] paid strictly by the pay-line, we would not have funded the grant that first helped identify the *BRCA1* gene". In addition to this, one of Richard Klausner's first initiatives when appointed director of NCI in the summer of 1995 was to institute an accelerated executive review process for initial applications that fall a little short of a fundable score. NCI funded 26 such petitions in the 1996 financial year, and has set aside money to fund about 35 this year.

Despite these policy changes, competition for funding for investigator-initiated grants is still fierce. Even so, there are actions an indi-

vidual can take that may give them an edge when submitting an application. Most young investigators, for example, underuse the scientific programme staff at NIH, in the belief that they are there to take care of their mentors, not them, says Robert A. Goldstein, NIAID's director of the division of allergy, immunology and transplantation. Most are willing to talk to colleagues at their own university or institution, where they may or may not get good advice, but "the first time they call us is after they get their first summary statement [the official summary of the study section review] that's got an unhappy score". NIAID tries to encourage young investigators to make contact earlier, he says.

Goldstein gets his message across through a variety of outreach efforts. His NIAID workshop, "Grantsmanship in Trying Times", will be held at the forthcoming AAAAI/AAI/CIS joint meeting. He says his opening remarks on such occasions tend to go like this: "We can't help you fix your science but... if you're doing good science, maybe we can help you write the grant application."

Weighing up whether to plump for an R29 first-investigator award (R29 grants are five-year awards designed to support newly independent investigators in the extramural research programme), or whether to apply for a larger RO1 grant (which are four years or

less), can be a tough decision for first-timers. It is another area where NIH programme officers may be able to help.

The review process and financial management policy in the various NIH institutes tend to be biased in favour of R29 applications. As a result, most institutes fund R29s to a slightly higher pay-line. "Technically, your odds would appear to be better when applying for an R29, if you think your score is going to be on the margin," says Thomas D. Williams, NIAID's budget officer. So, unless you are a hotshot PhD who has done a fabulous postdoc and has a handful of publications in top-flight journals, Goldstein advises newcomers to write an R29 and then, if the work goes well, to apply for an RO1 after two years. NIAID's booklet, *How to write an NIH grant*, which can be found on the Web, is a useful resource for first-time grant writers. □

Diane Gershon is assistant editor, new technology, Nature Medicine.

Molecular biology skills in demand in Europe

Claire O'Brien

"Any disease you care to mention, there's a T cell or a B cell or a macrophage in there somewhere," says postdoctoral immunologist Mark Smith from London's United Medical and Dental School. That is why his experience studying the cells that mount immune responses makes him "eminently employable" across many fields of medical science.

Even new graduates with an immunology BSc and one laboratory project under their belts can expect strong demand for their skills, says Jeremy Brock of the University of Glasgow. He organizes one of ten such dedicated BSc courses in the United Kingdom, and sees the demand — from industry and academic research — as being as strong now as when the course started in the mid-1970s. But, Brock adds, the skills required in this field have recently changed in an all-too-familiar way: "We've had to address criticisms that our courses didn't have enough molecular biology and cell signalling".

When Katherine Bodman-Smith was searching for her second immunology postdoctoral contract, she found that "a lot of them want molecular biology skills", which she could not offer. Instead, she landed a position using her specialist knowledge in the young field of glycobiology. Having studied how different carbohydrates on the surface of breast-tumour cells stimulate antibody responses, she moved sideways into parasitology to work on immunity to glycolipids expressed by *Leishmania* species, at the London School of Hygiene and Tropical Medicine. Bodman-Smith is funded on a three-year project grant from the Wellcome Trust,

whose current spending on grants that incorporate at least some immunology is £30 million (US\$48 million).

The UK Medical Research Council, like the Wellcome Trust, does not normally ring-fence research funds for specific topics, but does identify 'immunobiological manipulation' among the areas currently in need of trained staff and therefore particularly eligible for a 'training and career development award'. This research topic covers ways to boost insufficient immune responses to invading pathogens and, conversely, redirecting or reducing the response mounted against transplanted organs, or against self-antigens and allergens in those diseases where it results in irreversible tissue damage, such as insulin-dependent diabetes, rheumatoid arthritis and asthma.

The growing prevalence of such diseases, coupled with the global spread of both infectious and non-communicable diseases, all point to a need for clinically oriented immunology research alongside the molecular focus. More immunologists are seeking funds to take their work "out of the lab and into the clinic", says John Stephenson, scientific programme manager in infection and

immunity at the Wellcome Trust. He adds that this trend is encouraged by the many new diagnostic tools available for measuring immunological markers in serum. As a result, says Stephenson, "people studying the clinical aspects of immunology feature largely in our training and fellowship awards".

The commercial sector employs immunologists from graduate level upwards with specialisms ranging from molecules to whole animals. Alan Lamont, director of biology at Peptide Therapeutics in Cambridge, says he looks for "flexibility and enthusiasm" and team spirit in employees, "not a list of skills". Experience in a renowned laboratory, possibly abroad, and of working on projects from basic research through to a 'proof of principle' stage is also valued — attributes that would suit a candidate for biotechnology company jobs in any field.

But Lamont perceives a gap in the exploitation of immunology by drugs companies. "I get the impression that people controlling the purse strings in large pharmaceutical companies do not fully understand how the immune system works", he says. He cites the example of chronic viral disease research, where attention focuses on stopping the virus

replicating — for example, developing HIV protease and integrase inhibitors to slow the progress of AIDS — while the ability of the virus to 'hide' in a latent state is apparently forgotten. "If you can get a decent immune response going you have a much better chance of clearing the virus from the body — surely better in the long run!"

Another underexplored approach to combating viral disease, he adds, is to target the ways that viruses avoid the immune response, and allow the immune system to detect and clear the pathogen. For industrial research, "a real skill is... the ability to explain basic immunological systems" and companies will continue to need staff who understand the field, Lamont concludes. But "that can't be immunology for immunology's own sake". □

Claire O'Brien is a freelance science writer.

The British Society for Immunology Web site at <http://194.129.54.77/> is a useful site for immunologists, containing many links.

See also "Immunology resources on the World Wide Web" by J. Skipper and S. Man (*Molecular Medicine Today*, pp 320-323, August 1996) for a review of Web sites of interest to immunologists.

Bright prospects for British biotech

Molecular immunology is a rapidly progressing science. But the commercialization of researchers' findings is often slow. Scientists put a lot of faith in newly discovered 'key molecules', but few of these have been developed into effective treatments.

Hopes of effective cancer treatment have arisen from attempts by researchers to provoke an attack on a patient's immune system using genetically modified cells. The race is now on in this area, and especially in the United Kingdom there are several small companies developing gene therapies.

High-tech immunotherapies tend to be expensive and health-insurance companies may be unwilling to pay for them. So far in Germany, for example, industry has not embarked on developing techniques on a large scale.

In Germany, the number of vacancies for immunologists far outweighs the number of job-seeking scientists. Germany's main pharmaceutical companies prefer to conduct research in molecular immunology in countries such as the United States, Australia or Japan where there are fewer bureaucratic restrictions. Some small German biotech companies are optimistic about their future, but scientists in Germany are rarely employed for research and development, but rather for marketing,

consultation and service. Even though the environment for setting up a biotech company in France is less bureaucratic, the job market for scientists is limited as well. However, in the United Kingdom the prospects for immunologists are much brighter as the infrastructure for commercialization of biotechnology is more advanced and cooperation between universities and industry is more intense.

Companies in the United Kingdom often contract out research to universities. This provides essential contacts for university scientists, who could be taken on by the company if the project is successful. Small companies often take on staff in this way but the jobs are usually only short-term contracts.

The combination of cell biology and genetic engineering has led to promising new technologies and products from applied immunology. Start-up companies are taking their chances to offer their immunological know-how. An example of this new generation of companies in Germany is Freiburg-based CellGenix, founded in 1994 by a group of researchers and physicians in close contact with the university's medical centre. Progenitor blood cells, undifferentiated cells derived from the bone marrow, are isolated from the peripheral blood of patients before

undergoing high-dose chemotherapy. These cells are expanded by treatment with haematopoietic growth factors and can later be retransplanted. CellGenix executive Gregor Schulz expects that in Germany about 20,000 cancer patients a year could benefit.

But, while planning expansion of the company, Schulz complains about restrictions such as high taxes. He says there is no immediate prospect of a wave of start-up companies that might improve employment prospects in Germany. Most career opportunities for immunologists there will probably be provided in the research and development of diagnostic products. The blood banks make up a big potential market.

Serology, with its possibilities of standardization and automation, provides a growing market, for example in assessing the level of certain hormones, or detecting traces of an infarct or tumour. Vaccines make up another growing area. About 90 per cent of all vaccines are produced by 'classical' biochemical procedures. However, most new vaccines are recombinant proteins, a field in which US companies have a ten-year lead. And it is believed that the vaccines of the future will be intramuscular DNA injections, providing protection with a single application.

One thing is certain, the quest for new remedies will continue to provide jobs for immunologists. **Gabor Steigler**